

Aus dem Institut für Pathologie
der Heinrich-Heine-Universität Düsseldorf

Leiterin: Univ.-Prof. Dr. med. Irene Esposito
- Funktionsbereich Cytopathologie -

Erprobung der automatisierten DNA-Bildzytometrie
am Prostatakarzinom

Dissertation

zur Erlangung des Grades eines Doktors der Zahnmedizin
der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Mohammad Rezaie
2018

Als Inauguraldissertation gedruckt mit der Genehmigung der
Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

gez.:

Dekan: Univ.-Prof. Dr. med. Nikolaj Klöcker

Referent: Prof. Dr. med. Stefan Biesterfeld

Korreferent: PD Dr. med. Günter Niegisch

Meinen Eltern, Schwestern und meiner Frau gewidmet.

„Durch Wissen kommt der Mensch zur Menschlichkeit;

Fehlt Wissen ihm, gleicht er dem Tiere nur.

Thorheit ist Handeln in Unwissenheit,

und Thorheit findet nie der Wahrheit Spur.“ (Hafez)

Zusammenfassung

Die DNA-Bildzytometrie repräsentiert ein TV-bildanalytisches Messverfahren, mit welchem an Feulgen-gefärbten zytologischen Präparationen das DNA-bildzytometrische Profil von Zellpopulationen bestimmt wird. Dabei wird das Prinzip genutzt, dass über die stöchiometrische Feulgen-Färbung der DNA-Gehalt von Zellkernen im sichtbaren Licht durch die Messung ihrer integrierten optischen Dichte (IOD) im Durchlichtverfahren quantifiziert werden kann. Nach Vergleich mit den IOD-Werten einer Referenzzellpopulation kann dann entschieden werden, ob eine bestimmte Analysezellpopulation, also z. B. tumorverdächtige Zellen, über eine numerische Chromosomenaberration verfügen oder nicht. Die Methode wird am Schwerpunkt Cytopathologie der Heinrich Heine-Universität Düsseldorf unter anderem an Zellvereinzelungspräparaten von Stanzbiopsien histologisch gesicherter Prostatakarzinome eingesetzt und dient hier im Wesentlichen dem Zweck, dem Patienten und seinen behandelnden Ärzten ergänzende Informationen über die Möglichkeit einer *Active Surveillance*-Strategie an die Hand zu geben.

Klassischerweise werden DNA-bildzytometrische Messungen durchgeführt, indem ein *Operator* per Mausklick die zu bewertenden Zellkerne auswählt und so unter ständiger Kontrolle aller Vorgänge an einer begrenzten Stichprobe von ca. 300 Analysezellen pro Messung das DNA-bildzytometrische Profil ermittelt. Dieser „manuellen Messung“ könnte, sofern sie sich aus messtechnischer und klinischer Sicht bewährt, eine „automatisierte Messung“ gegenüber gestellt werden. Bei dieser neu entwickelten und hier erstmals an Zellvereinzelungspräparaten erprobten Methode läuft der eigentliche Messvorgang nach Einrichtung des Gerätes vollautomatisiert ab, wobei eine wesentlich größere Zellzahl als bei der manuellen Messung erfasst wird. Dem *Operator* obliegt dann nach Ende der Messung anhand der Bildgalerie, in der alle Objekte abgelegt werden, die Kontrolle darüber, ob die gemessenen Objekte wirklich Zellen entsprechen (in Abgrenzung gegenüber Artefakten) und ob sie richtig als Referenzzellen oder Analysezellen zugeordnet sind. Die eigentliche klinische Interpretation der Messung wird bei beiden Messverfahren fachärztlich durch

einen Pathologen vorgenommen, der auch den schriftlichen Befund erstellt und ausfertigt.

In der hier vorliegenden Studie wurden 27 Fälle von Prostatakarzinomen zweimal automatisiert mit einem TV-Bildanalyse-System MotiCyte-DNA (Fa. Motic) nachgemessen, für die in der klinischen Routine bereits ein manuelles DNA-bildzytometrisches Profil erstellt und als Typ B (n = 12) bzw. Typ C (n = 15) nach Tribukait (1993) klassifiziert worden war. Als Messparameter wurden jeweils vor und nach Bearbeitung der Messobjekte durch den *Operator* a) für die Referenzzellen ihre Zahl, der Variationskoeffizient (CV) ihrer IOD sowie der Korrelationskoeffizient (CC) zwischen der mittleren IOD und der Kernfläche ausgewertet sowie b) für die Analysezellen ebenfalls ihre Zahl, ferner die Höhe des maximalen Messwertes (C-Max) und die Lage des ersten bzw. zweiten Peaks der Messwertverteilung.

Als methodische Ergebnisse lässt sich festhalten, dass sich die Zahl der akzeptierten Objekte zwar stark unterschied, sich ansonsten aber die Ergebnisse der drei Messreihen im Wesentlichen gut korrelieren ließen. Einige Schwächen der Softwareversion, die 2014/15 bei Durchführung der Messungen zur Verfügung stand, betrafen die Fehlzuordnung von sich überschneidenden Zellkernen oder von eindeutig defokussierten Objekten zu den Analysezellen statt zu den Artefakten. Außerdem war das Vorgeben einer gewissen Fokusebene durch das Setzen sogenannter „Fokuspunkte“, innerhalb derer das Meßsystem den Autofokus anwenden sollte, fehlerbehaftet. Diese Schwächen kosteten relativ viel Zeit, so dass mit unserer Softwareversion das Ziel einer Zeitersparnis verglichen mit der manuellen DNA-bildzytometrischen Messung nicht erreicht werden konnte.

Bezogen auf die klinische Vergleichbarkeit lässt sich sagen, dass beide automatisierten Messungen jeweils zum identischen Ergebnis kamen, d.h. entweder zweimal Typ B oder zweimal Typ C nach Tribukait ermittelt wurde. Im Vergleich zur manuellen Messung stimmten 22 der 27 Zuordnungen überein (= 81,5%). Drei manuell als Typ B zugeordnete Fälle wurden automatisiert als Typ C eingestuft, in zwei Fällen war es umgekehrt.

Inzwischen wurden, auch basierend auf unseren Erfahrungen, Verbesserungen an der Software und Modifikationen an der Geräteeinstellung vorgenommen, so dass die automatisierte DNA-Bildzytometrie nunmehr an mehreren Standorten, darunter auch am Schwerpunkt Cytopathologie der Heinrich Heine-Universität Düsseldorf, in der klinischen Diagnostik angewandt werden kann.

Abstract

The DNA image cytometry represents a TV-image-analyzing-measurement method, which gives information about the DNA image cytometrical profile of a Feulgen stained cytological preparation.

This method quantifies the DNA content of the nucleus through the measurement of the integrated optical density (IOD) of the Feulgen stained DNA content. After comparing their IOD value with the reference IOD value a statement can be made about the numeric chromosome-aberration of the analyzing cells. This method is used in the center of Cytopathology of the Heinrich-Heine-University. Amongst other things it is used for cell separation preparation of core biopsy of prostate cancer and is used to give additional information about the "Active Surveillance" treatment for the treating physician.

These measurements are usually done by an operator, who selects about 300 analyzing cells manually. This part can be automated in the future if the machine proves itself.

This newly developed method, which is tested on core biopsy for the first time, runs the whole measurement automatically after initial setup and collects far more cells than manually performed measurements. Once the measurement is done, the operator will have to check the image gallery for falsely classified cells, which are placed in the wrong category. For both methods a physician will have to interpret the results of the measurement and give a diagnosis.

In this study 27 cases of prostate cancer, which were previously routinely measured manually, were measured automatically twice with the TV-image-analyzing-system MotiCyte-DNA (Motic Co.). There were 12 Type B and 15 Type C classifications after Tribukait (1993). The measurement parameters were a) quantity of reference cells, the variationcoefficient (CV) of their IOD and the correlationcoefficient (CC) between the average IOD and its nucleus surface; b) quantity of the analyzing cells, the highest measured DNA content (C-Max), and the positioning of the first and second peak.

The results show that the quantity of the selected cells varies dramatically, but the interpretation of the measurement mostly correlates. Some weak points of the software-version from 2014/2015 were the falsely classification of the cells and the classification of some obviously defocused cells to the analyzing cell population, instead of the artifact population. Furthermore, there were some errors in the system during the setup of the focus level. These errors were time consuming, compared manually measured method

As for the clinical interpretation the automated measurement always came to the same conclusion, both times either Type B or Type C. Compared to the manual method 22 out of 27 (81,5%) measurements correlated. Three manually measured Type Bs were diagnosed Type C and two manually measured Type Cs were diagnosed Type B.

Meanwhile based on our experience, improvements on the software and some modification to the device settings were made, so currently the automated DNA image cytometry is used in many departments at the Heinrich-Heine-University.

Abkürzungsverzeichnis

°C	Grad Celsius
µm	Mikrometer
1D	Unterprogramm des BMDP-Statistikprogramms
2D	Unterprogramm des BMDP-Statistikprogramms
3D	Unterprogramm des BMDP-Statistikprogramms
4F	Unterprogramm des BMDP-Statistikprogramms
5cEE	<i>5c exceeding events</i> , Zellen mit einem DNA-Gehalt > 5c
9cEE	<i>9c exceeding events</i> , Zellen mit einem DNA-Gehalt > 9c
Abb.	Abbildung
AS	<i>active surveillance</i> , aktive Überwachung
AWMF	Arbeitsgemeinschaft Wissenschaftlich-Medizinischer Fachgesellschaften
BMDP	Name eines verwendeten Statistikprogramms
BPS	Bundesverband Prostatakrebs Selbsthilfe
BRCA-1-Gen	<i>breast-cancer-1-Gen</i>
BRCA-2-Gen	<i>breast-cancer-2-Gen</i>
bzw.	beziehungsweise
c	Einheit für den DNA-Gehalt, bei normalen Zellen 2c
CC	<i>coefficient of correlation</i> , Korrelationskoeffizient; hier: zwischen IOD der Eichzellkerne und ihrer Kernfläche
C-Max	Maximaler DNA-Wert einer Zellpopulation
CML	chronisch-myeloische Leukämie
cT	klinische Beurteilung des Primärtumors in der TNM-Klassifikation
CV	<i>coefficient of variation</i> , Variationskoeffizient
dest.	<i>destillata</i> , destilliert
DNA	<i>desoxyribonucleic acid</i> , Desoxyribonukleinsäure
DRU	digitale rektale Untersuchung

ESACP	<i>European Society for Analytical and Cellular Pathology</i>
g	Einheit für die Erdbeschleunigung
G ₀ /G ₁ -Fraktion	ruhende Zellpopulation, DNA-Gehalt der Zellen bei 2c
G ₂ /M-Phase	Zellzyklusphase nach der Proteinbiosynthese, DNA-Gehalt der Zellen bei 4c
G ₂ /M'-Phase	Zellzyklusphase nach zweimaligem Durchlaufen der Proteinbiosynthese ohne zwischenzeitliche Zellteilung, DNA-Gehalt der Zellen bei 8c
G ₂ /M'''-Phase	Zellzyklusphase nach viermaligem Durchlaufen der Proteinbiosynthese ohne zwischenzeitliche Zellteilung, DNA-Gehalt der Zellen bei 32c
GHz	Gigahertz
GS	Gleason-Score
H ₀	Nullhypothese
HCl	Hydrochlorid bzw. Bezeichnung für Salzsäure
IOD	integrierte optische Dichte
Kap.	Kapitel
M	Mittelwert
Mess.	Messung
min.	Minute
ml	Milliliter
MTA	Medizinisch-Technische(r) Assistent(-in)
N	Einheit für Normalität in der Chemie, z.B. 5-normale Salzsäure = 5N HCl
ng	Nanogramm
NICE	<i>National Institute for Health and Clinical Excellence</i>
p	Wahrscheinlichkeit
PSA	Prostata-spezifisches Antigen
r	Korrelationskoeffizient
R	Beurteilung des Resektionsrandes in der TNM-Klassifikation
RKI	Robert Koch-Institut

RWTH	Rheinisch-Westfälische Technische Hochschule (Aachen)
S3	höchstmögliches Qualitätsmerkmal einer Leitlinie der AWMF (Stufen: S1, S2k, S2e, S3)
SD	<i>standard deviation</i> , Standardabweichung
SO ₂	Schwefeldioxid, hier in wässriger Lösung als SO ₂ -Wasser verwendet
SOP	<i>standard operation procedures</i> , schriftlich hinterlegte Standardvorgehensweisen im Labor
T	Beurteilung des Primärtumors in der TNM-Klassifikation
TNM	Beurteilungssystem für maligne Tumoren nach Ausdehnung des Primärtumors (T), ggfls. befallenen Lymphknoten (N) und eventuellen Fernmetastasen (M) gemäß den Vorgaben der UICC
Tab.	Tabelle
TV	<i>television</i> , hier: TV-Bildanalyse, d.h. Bildanalyse am von einer Kamera auf einen Monitor übertragenen Bild
UICC	<i>Union of International Cancer Control</i> , eigentlich <i>Union internationale contre le cancer</i>
UKD	Universitätsklinikum Düsseldorf
vgl.	vergleiche
W	Watt
x	-fach; z.B. 40-fache Vergrößerung (40x)
x	Bezeichnung der Abszisse im Koordinatensystem
y	Bezeichnung der Ordinate im Koordinatensystem

Inhaltsverzeichnis

1	Einleitung	1
1.1	Das Prostatakarzinom - eine häufige Tumorentität	1
1.2	Stellenwert der Prostatastanzbiopsie in der Diagnostik des Prostatakarzinoms	4
1.3	DNA-Bildzytometrie als ergänzende Methode in der Malignitätsabschätzung des Prostatakarzinoms	5
1.4	Methodisches Prinzip der DNA-Bildzytometrie	7
1.5	Anwendung der DNA-Bildzytometrie beim Prostatakarzinom	8
1.6	Automatisierte DNA-Bildzytometrie	11
2	Ziel der Arbeit	15
3	Material und Methoden	17
3.1	Auswahl der Messpräparate	17
3.2	Zellvereinzellung und Präparateherstellung	18
3.3	Feulgen-Färbung	19
3.4	Manuelle DNA-Bildzytometrie	21
3.5	Automatisierte DNA-Bildzytometrie	24
3.6	Aufbau der Messreihe	33
3.7	Untersuchte Messparamater	34
3.8	Statistik	38
4	Ergebnisse	39
4.1	Patientenkollektiv	39
4.2	Methodische Untersuchungen an den Referenzzellen	39
4.2.1	Zahl der Referenzzellen	39

4.2.2	Variationskoeffizient (CV) der integrierten optischen Dichte (IOD) der Referenzzellen	43
4.2.3	Korrelationskoeffizient (CC) zwischen der mittleren IOD und der Kernfläche der Referenzzellen	47
4.3	Methodische Untersuchungen an den Analysezellen	51
4.3.1	Zahl der abnormen Zellen	51
4.3.2	Höhe des maximalen Messwertes (C-Max)	56
4.3.3	Lage des ersten Peak	60
4.3.3.1	Tribukait B-Fälle	61
4.3.3.2	Tribukait C-Fälle	64
4.3.4	Lage des zweiten Peak	69
4.3.4.1	Tribukait B-Fälle	69
4.3.4.2	Tribukait C-Fälle	73
4.4	Klinische Datenauswertung	77
4.4.1	Altersverteilung der Patienten	77
4.4.2	Gegenüberstellung der Ergebnisse von manueller und automatisierter DNA-Bildzytometrie	77
4.4.3	Diskrepant beurteilte Fälle	78
5	Diskussion	81
5.1	Das Prostatakarzinom - ein „Sonderfall“ in der klinischen Tumormedizin	81
5.2	Grundzüge der „Active Surveillance“-Strategie	85
5.3	DNA-Bildzytometrie als ergänzendes Auswahlkriterium für die Wahl einer „Active Surveillance“-Strategie	88
5.4	Automatisierte DNA-Bildzytometrie - Möglichkeiten und Grenzen	93
6	Literaturverzeichnis	99

1 Einleitung

1.1 Das Prostatakarzinom - eine häufige Tumorentität

Das Prostatakarzinom stellt mit 25,3% aller Krebsneuerkrankungen die häufigste Krebsart bei Männern in Deutschland (Abb. 1). Im Jahr 2012 gab es 63.710 Neuerkrankungen (RKI 2015¹). Außerdem ist das Prostatakarzinom nach Lungen- und Darmkrebs die dritthäufigste tödlich verlaufende Krebserkrankung bei Männern, in Zahlen sind es etwa 12.000 Patienten jährlich in Deutschland (Beske et al. 2007, Rohde et al. 2007).

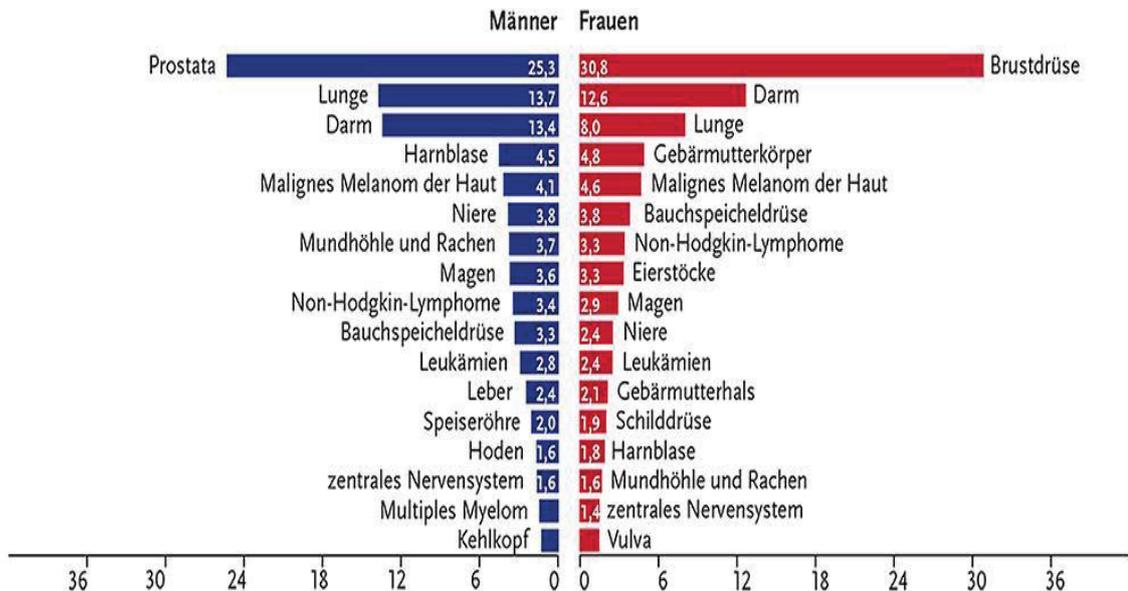


Abb. 1: Aktuelle Krebsstatistik zur Inzidenz von malignen Tumoren in Deutschland 2012 (RKI 2015). Angaben geschlechtsgetrennt in Prozent.

Eindeutige kausale Ursachen für die Tumorentstehung sind nicht bekannt. Es wurden allerdings Risikofaktoren identifiziert, bei deren Vorliegen Prostatakarzinome häufiger auftreten.

¹ Die Literaturangaben mit der Abkürzung „RKI“ beziehen sich im Folgenden stets auf die gemeinsam vom Robert Koch-Institut und der Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. gemeinsam herausgegebenen Broschüren „Krebs in Deutschland“ bestimmter Zeitabschnitte und bestimmter Auflagen; siehe Literaturverzeichnis.

Der wichtigste Risikofaktor ist das Alter des Patienten. Bei den 30- bis 34-Jährigen erkrankt nur 1/100.000 und Jahr, der Höhepunkt der Inzidenzrate liegt in der Altersgruppe 65- bis 69-Jährigen mit ca. 750/100.000 und Jahr (Eaton et al. 1999, RKI 2015) (Abb. 2). Es handelt sich somit um eine Erkrankung vorwiegend älterer Männer.

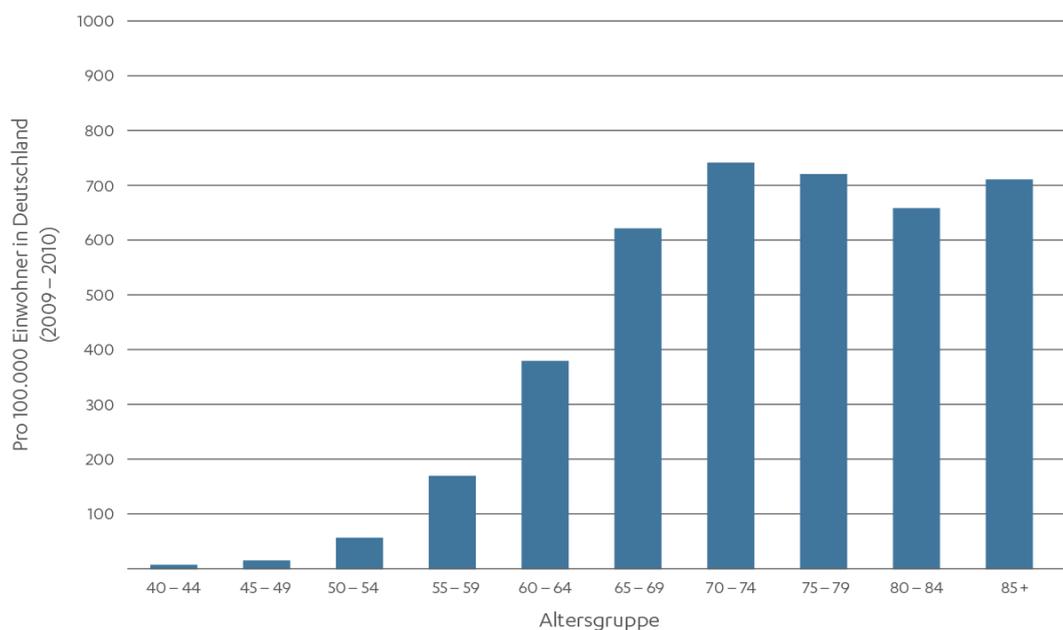


Abb. 2: Altersbezogene Inzidenz beim Prostatakarzinom in Deutschland 2009-2010 (RKI 2015)

Ein weiterer wichtiger Risikofaktor, der das Erkrankungsrisiko erhöht, ist die genetische Disposition. Für einen Patienten, dessen Vater oder Bruder an einem Prostatakarzinom erkrankt ist, wird in der aktuellen Leitlinie „*Interdisziplinäre Leitlinie der Qualität S3 zur Früherkennung, Diagnose und Therapie der verschiedenen Stadien des Prostatakarzinoms*“, Langversion 4.0 von Dezember 2016, AWMF-Register-Nummer 043/022OL, im Folgenden kurz als „S3-Leitlinie“ bezeichnet (Leitlinienprogramm Onkologie 2016), ein zweifach erhöhtes Erkrankungsrisiko angegeben. Bei zwei bzw. drei betroffenen erstgradigen Verwandten werden ein fast fünffaches bzw. fast elfaches relatives Risiko genannt (Hemminki und Czene 2002, Bruner et al. 2003, Paul et al. 2008). Zytogenetische Untersuchungsergebnisse zur Identifikation von genetischen Risikoprofilen, z.B. bei Keimbahnmutationen im BRCA-1- oder BRCA-2-Gen, wie sie aktuell von Cheng et al. (2017) zusammengefasst wurden, mündeten in Empfehlungen engerer Kriterien zur Durchführung von

Prostatastanzbiopsien, haben aber noch keinen Eingang in die aktuelle S3-Leitlinien gefunden.

Bei abgelaufener Prostatitis wird ein ca. 1,5-faches Erkrankungsrisiko gesehen (Leitlinienprogramm Onkologie 2016). Weitere allgemein anerkannte Risikofaktoren existieren nicht; die Literaturdiskussion in der aktuellen S3-Leitlinie kann man so zusammenfassen, dass Studien zu Einflüssen von Ernährungsweise, Z.n. Vasektomie, Diabetes mellitus oder Adipositas bisher widersprüchliche oder unklare Ergebnisse zeigten und z. Zt. keine verwertbaren Hinweise auf eine Inzidenzerhöhung von Prostatakarzinomen erlauben.

Bei Verdacht auf ein Prostatakarzinom werden folgende primäre diagnostische Maßnahmen gemäß der aktuellen S3-Leitlinie von 2016 empfohlen:

1. digital-rektale Untersuchung,
2. Bestimmung des Serumwertes für das Prostata-Spezifische Antigen (PSA),
3. transrektale Ultraschalluntersuchung,
4. Prostatastanzbiopsie

Zu allen vier Methoden ist zu sagen, dass sie nicht in jedem Fall eine eindeutige Tumordiagnose bzw. die Klärung eines Tumorverdacht ermöglichen, da stets Differenzialdiagnosen zu bestimmten Untersuchungsergebnissen zu bedenken sind. So werden nur ca. 2%-5% der Prostatakarzinome durch die digital-rektale Untersuchung (Lavoipierre et al. 1998, Hoogendam et al. 1999, RKI 2012) erkannt. Ähnlich ist auch das PSA kein besonders harter Marker, da es einen recht breiten Überlappungsbereich zwischen den PSA-Werten von prostatistischen Erkrankungen verschiedener Dignität gibt. So kann ein erhöhter PSA-Wert auch durch Eingriffe am unterem Harntrakt, durch Harnwegsinfekte oder Leberfunktionsstörungen bedingt sein (Fornara et al. 2004, Rohde et al. 2007), so dass eine relativ breite „Grauzone“ von Werten, bei denen der PSA-Wert nicht eindeutig in eine spezielle Richtung deutet, existiert. In der wissenschaftlichen Literatur werden bei positivem digitalen Tastbefund folgende Häufigkeiten einer Diagnose eines Prostatakarzinoms in Abhängigkeit vom PSA-Wert angegeben (Cooner et al. 1990, Paul et al. 2008):

- < 4 ng/ml	11,7%,
- 4,1 - 10 ng/ml	42,6%
- > 10 ng/ml	76%

Für PSA-Werte zwischen 4,1 und 10 ng/ml wurde der positive prädiktive Wert für ein Prostatakarzinom mit 25%-35% veranschlagt, für Werte > 10 ng/ml mit 50%-80% (Paul et al. 2008).

Das „härteste“ Diagnostikum zur Erkennung eines Prostatakarzinoms ist die Prostatastanzbiopsie, die heute in Deutschland als morphologisches Standardverfahren etabliert ist. Die in Nordeuropa durchaus nicht unübliche Prostatazytologie war in Deutschland nie weit verbreitet und wird kaum noch durchgeführt. Zu Kriterien für die Stanzung und zu der Methodik der Aufarbeitung und diagnostischen Interpretation sei auf die aktuelle S3-Leitlinie verwiesen.

1.2 Stellenwert der Prostatastanzbiopsie in der Diagnostik des Prostatakarzinoms

Deutsche und internationale Leitlinien empfehlen zur Sicherung einer Verdachtsdiagnose auf Prostatakarzinom, insbesondere bei einem erhöhtem PSA-Wert oder einem auffälligen Tast- bzw. Ultraschallbefund, eine Ultraschallkontrollierte Stanzbiopsie durchzuführen. Hierzu sollen vorzugsweise aus dem rechten und aus dem linken Prostatalappen jeweils sechs anatomisch definierte Biopsien entnommen werden und einzeln mit topographischer Zuordnung zur histologischen Beurteilung eingesandt werden. Die Untersuchung erfolgt unter Lokalanästhesie und Antibiotikaschutz.

Die Präparate werden in der Pathologie standardisiert aufgearbeitet und histologisch beurteilt. Im Falle eines Tumorbefundes erfolgt gemäß der S3-Leitlinie pro Stanze die Angabe des prozentualen Gewebsbefalls durch den Tumor und die Gradierung nach dem international anerkannten modifizierten Gleason-Score (Epstein et al. 2005).

In die Therapieplanung gehen neben klinischen Größen (Alter, Allgemeinzustand des Patienten, limitierende Vorerkrankungen) und den Ergebnissen der bildgebenden Ausbreitungsdiagnostik mit Schnittbilddiagnostik und Szintigraphie auch die Zahl der betroffenen Stenzen und der Gleason-Score als morphologische Größen ein. Mögliche Therapieoptionen reichen vom einfachen Abwarten („*Watchful Waiting*“) über bewusstes nicht direktes therapeutisches Handeln mit engmaschigen Kontrollen („aktives Abwarten“ bzw. „*Active Surveillance*“, AS) bis hin zu hormonellen bzw. strahlentherapeutischen Ansätzen bzw. zur operativen Therapie in Form einer radikalen Prostatektomie, ggfls. mit Lymphadenektomie.

Der Gleason-Score ist dabei in der wissenschaftlichen Literatur nicht unumstritten, da seine Reproduzierbarkeit zwischen zwei oder mehreren Untersuchern nicht besonders hoch ist (siehe auch Diskussion). Außerdem ist er als einziges Gradierungssystem bei epidemiologisch wichtigen Tumoren rein auf das Wachstumsmuster der Tumorzellen fokussiert, während die Zellmorphologie, ausgedrückt durch Kernpolymorphie und Mitoserate, keine Berücksichtigung findet. Dieses ist anders gerade auch bei den anderen drei führenden Organkarzinomen (Mamma, Lunge, Darm).

1.3 DNA-Bildzytometrie als ergänzende Methode in der Malignitätsabschätzung des Prostatakarzinoms

Als eine ergänzende Methode für eine Gradierung steht beim Prostatakarzinom die DNA-Bildzytometrie zur Verfügung. Diese Anwendung ist in Deutschland insbesondere von den Arbeitsgruppen um Böcking (Aachen, später Düsseldorf) und Haroske (Dresden) gefördert worden, insbesondere hier auch durch die Entwicklung von zur Durchführung geeigneter TV-Bildanalysesysteme, die in Kooperation mit Technischen Universitäten und optischen Herstellern wie Zeiss (Oberkochen), Hund (Wetzlar) oder Leica (Wetzlar) bzw. aktuell mit der chinesischen Firma Motic (Xianmin, China und Wetzlar) umgesetzt worden sind.

Die diagnostischen Möglichkeiten der DNA-Bildzytometrie sind in diversen Publikationen belegt, insbesondere auch für glanduläre Läsionen an der

Schleimhaut des Endometriums (Biesterfeld et al. 2001a), der Endozervix (Biesterfeld et al. 2001b) oder der Gallenwege (Krishnamurthy et al. 2001, Osterheld et al. 2005, Biesterfeld und Deacu 2009) sowie bei Borderline-Läsionen des Ovars (Fleischer 2017) und Dysplasien der Bronchialschleimhaut (Auffermann und Böcking 1985), wobei die aktuellen Vorgaben an Messpräzision und Interpretation (siehe Kapitel 1.4) erfüllt waren. Auch an Ergusspräparaten, die verschiedene glanduläre Neoplasien enthalten kann, war die Methode sinnvoll anwendbar (Motherby et al. 1998).

Genutzt wird sie aber nur an wenigen diagnostischen Einrichtungen, darunter insbesondere auch am Schwerpunkt Cytopathologie der Heinrich-Heine-Universität Düsseldorf. Dort werden pro Jahr ca. 800 Messungen pro Jahr an diagnostisch unklaren Fällen verschiedener Organe durchgeführt, um zu einer abschließenden diagnostischen Entscheidung zu gelangen (Biesterfeld et al. 2013): Wichtigste Indikation zur Durchführung einer Messung ist die Abklärung von in der konventionellen Zytologie zweifelhaften oder dringend verdächtigen Veränderungen, z. B. bei unklaren Veränderungen in Lungenpunktaten, Gallengangabstrichen oder Punktaten des Pankreas. Auch die Untersuchung von Prostatastanzbiopsien, vor allem bei Konstellationen, in denen eine *Active Surveillance*-Strategie diskutiert wird, stellt eine wesentliche Indikation dar, die auch Grundlage für die Durchführung der hier vorgestellten Studie ist.

Das Prostatakarzinom ist derzeit der einzige Organtumor, bei dem die DNA-Bildzytometrie zur Gradierung eingesetzt wird. Diese Nutzung geht insbesondere auf die unten näher erläuterte grundlegende Arbeit von Tribukait (1993) zurück, der für ein großes Kollektiv von Patienten mit Prostatakarzinomen eine prognostische Relevanz der Methode belegen konnte. Alle diese Patienten waren unbehandelt, weswegen die Arbeit nach wie vor als Datenbasis dafür verwendet wird, aus dem DNA-zytometrischen Profil eines Prostatakarzinoms Rückschlüsse auf die Chancen und Risiken einer therapeutisch abwartenden Haltung zu ziehen. Danach sind eine Reihe anderer Studien zur DNA-Bildzytometrie beim Prostatakarzinom veröffentlicht worden, die mit einer prognostischen Relevanz auch behandelter Patientengruppen belegt haben. Hierzu soll auf einen aktuellen Review aus dem Jahr 2014 verwiesen werden, in welchem die Ergebnisse von 67 prognostischen DNA-zytometrischen Studien Eingang gefunden haben (Böcking et al. 2014).

1.4 Methodisches Prinzip der DNA-Bildzytometrie

Die DNA-Bildzytometrie ist eine Methode, bei der TV-bildanalytisch eine mikrodensitometrische Quantifizierung des DNA-Gehaltes von klassischerweise ca. 250 - 300 morphologisch auffälligen Zellen und von ca. 30 normalen Eichzellen vorgenommen wird. Sie kann an allen zytologischen Aufbereitungen (Zytozentrifugaten, Sedimentausstrichen, Abstrichen) erfolgen, ebenso an nach enzymatischer Zellvereinzelung hergestellten Ausstrichen von Zellen aus Paraffinblöcken. Voraussetzung ist eine spezifische und stöchiometrische Darstellung der DNA, die über eine Feulgen-Färbung realisiert wird. Vorgaben an die Messmethodik wurden im Rahmen eines internationalen Konsensus der ESACP („European Society for Analytical and Cellular Pathology“) weitgehend vereinheitlicht (Böcking et al. 1995, Giroud et al. 1998, Haroske et al. 1998, Haroske et al. 2001); dies gilt auch für wesentliche Bestandteile der Auswertalgorithmen. Ziel der DNA-Bildzytometrie ist die Abklärung einer etwaigen numerischen DNA-Aneuploidie, welche man bei Darstellung einer aneuploiden DNA-Stammlinie als Nachweis einer chromosomalen abnormen klonalen Zellpopulation oder von Einzelzellaneuploidie diagnostiziert. In Abbildung 3 ist ein Ausschnitt aus einem Messvorgang dokumentiert.

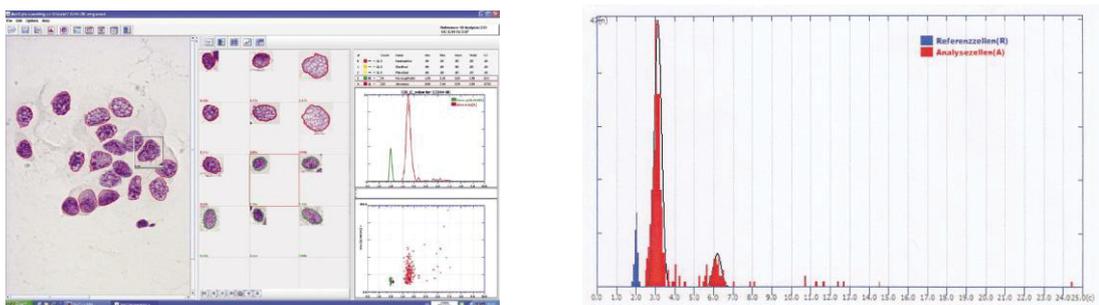


Abb. 3: Links: Monitorarstellung eines aktuellen Messvorganges mit einem modernen DNA-bildzytometrischen Meßsystem (MotiCyteDNA) mit Darstellung des mikroskopischen Livebildes, der Bildgalerie der letzten 12 erfassten Objekte sowie Diagrammdarstellungen der DNA-Messwertverteilung (oben) und der Beziehung zwischen DNA-Messwerten und Kernflächen (unten). Rechts: Ergebnisausdruck eines DNA-aneuploiden Falles mit einer DNA-Stammlinie bei 3,05c und einem Verdopplungsgipfel bei 6,25c (Biesterfeld 2014)

Der erwähnte Begriff der „numerischen DNA-Aneuploidie“ hebt dabei rein auf die Gesamtmenge chromosomaler DNA ab und umfasst nicht Aspekte der „chromosomalen DNA-Aneuploidie“, die sich mit strukturellen Auffälligkeiten an Chromosomen beschäftigt. Da in aller Regel aber in malignen Tumoren numerische DNA-Aneuploidie vorliegt, die von einem bestimmten Ausmaß her aus messbar sein wird, werden fast alle Tumorentitäten durch DNA-Bildzytometrie identifizierbar. Eine wichtige Ausnahme stellt die chronisch-myeloische Leukämie (CML) dar, die eine rein chromosomale Aneuploidie, gekennzeichnet durch die sogenannte Philadelphia-Translokation, aufweist.

1.5 Anwendung der DNA-Bildzytometrie bei Prostatakarzinomen

Wie erwähnt dient die DNA-Bildzytometrie beim Prostatakarzinom nicht der Diagnostik, also der Entscheidung, ob überhaupt ein Karzinom vorliegt, sondern der Abschätzung seiner biologischen Aggressivität, dem *Grading*.

Diese Anwendung geht zurück auf eine Studie zum Langzeitüberleben von knapp 300 Patienten mit unbehandeltem Prostatakarzinom (Tribukait 1993). Beim DNA-zytometrischen Profil der Tumoren wurden in der Originalarbeit drei Gruppen (A, B und C) unterschieden. In Gruppe A wurden Tumoren eingeordnet, die ausschließlich eine diploide DNA-Stammlinie aufwiesen, in Gruppe B Tumoren mit einer diploiden und einer tetraploiden DNA-Stammlinie und in Gruppe C Tumoren mit aneuploider DNA-Verteilung in Form einer oder mehrerer aneuploider DNA-Stammlinien und / oder Einzelzellaneuploidie. Die Kaplan-Meier-Überlebenskurven belegten einen deutlichen Unterschied in der Überlebenswahrscheinlichkeit der Patienten. Später wurde in Anlehnung an die Auer-Klassifikation für Mammakarzinome (Auer et al. 1980) die Gruppe C in zwei Subgruppen (eine aneuploide DNA-Stammlinie [wiederum mit C bezeichnet] vs. ungeordnet aneuploide DNA-Verteilung, sogenannte *Manhattan skyline*, [D]) untergliedert, auf die auch in der Modifizierung der Originalabbildung Bezug genommen wird (Abb. 4). Während der Überlebensunterschied zwischen den Gruppen A und B nur einen nicht signifikanten Trend widerspiegelte, war der Überlebensunterschied zwischen den Gruppen B und C nicht nur graphisch eindeutig, sondern auch statistisch signifikant ($p < 0,0001$).

Auf diese Arbeit geht auch die heutige, am Universitätsklinikum Düsseldorf verwendete DNA-bildzytometrische Klassifikation von Prostatakarzinomen zurück (Böcking und Dietz 2013) (Tab. 1). Typische DNA-Histogramme sind in Abbildung 5 wiedergegeben.

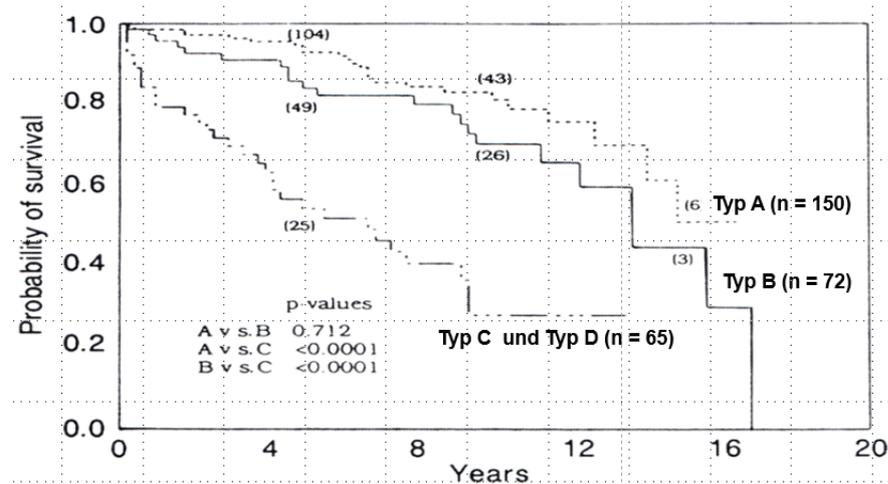


Abb. 4: Kaplan-Meier-Überlebenskurven für das Langzeitüberleben von 287 Fällen unbehandelter Prostatakarzinome nach dem DNA-bildzytometrischen Profil des Tumors (Biesterfeld 2014, modifiziert nach Tribukait 1993).

DNA-Gradierung	Definition	Prognose
1 Peridiploid (Typ A)	Eine DNA-Stammlinie bei $2c \pm 10\%$	Gut
2 Peritetraploid (Typ B)	Eine DNA-Stammlinie bei $2c \pm 10\%$, zweite Stammlinie bei $4c \pm 10\%$ oder mind. eine Zelle $>4,4c$	relativ gut
3 X-Ploid (Typ C)	Eine zusätzliche DNA-Stammlinie außerhalb $2c \pm 10\%$ und $4c \pm 10\%$ oder mind. eine Zelle $>9c$	schlechter
4 Multiploid (Typ D)	Mehrere DNA-Stammlinien außerhalb $2c \pm 10\%$ oder $4c \pm 10\%$	schlecht

Tab. 1: Einteilung und prognostische Bedeutung der DNA-bildzytometrischen Gradierung beim Prostatakarzinom (Böcking und Dietz 2013).

Typische DNA-Histogramme	Diagnose DNA-Grad vs. Gleason-Score	Prognose
	Peridiploide DNA-Verteilung DNA-Grad 1 (Typ A) entspricht etwa GS ≤ 6	Sehr gut Aktive Überwachung bei Mikrokarzinomen Befund-Häufigkeit in Stanzbiopsien: ca. 27% der Fälle
	Peritetraploide DNA-Verteilung DNA-Grad 2 (Typ B) entspricht etwa GS 7	Überwiegend gut Für ältere Patienten wie bei Grad 1 (Typ A) Befund-Häufigkeit in Stanzbiopsien: ca. 52% der Fälle
	X-ploide DNA- Verteilung DNA-Grad 3 (Typ C) entspricht etwa GS 8	Unbehandelt nicht mehr so gut Behandlung wie bei Prostata-Karzinom mit Gleason ≥ 8 Befund-Häufigkeit in Stanzbiopsien: ca. 11% der Fälle
	Multiploide DNA- Verteilung DNA-Grad 4 (Typ D) entspricht etwa GS 9 - 10	Unbehandelt nicht gut Behandlung wie bei Prostata-Karzinom mit Gleason ≥ 9 Befund-Häufigkeit in Stanzbiopsien: ca. 10% der Fälle

Abb. 5: Typische DNA-Histogramme beim Prostatakarzinom mit üblicher Korrelation zu Gleason-Score (GS), Prognose und Häufigkeit in Stanzbiopsien (Böcking und Dietz, 2013, Biesterfeld 2014). Referenzzellen blau, Analysezellen rot.

Dass diese Histogrammklassifikation über eine hohe Reproduzierbarkeit unter mehreren Untersuchern verfügt, konnte in wissenschaftlichen Arbeiten der Heinrich-Heine-Universität gezeigt werden; bei der Re-Evaluation von 103 bzw. 456 DNA-Histogrammen von Messungen an Prostatakarzinomen ergaben sich Übereinstimmungswerte von 93% bzw. 90.2% (Tils 2016).

Bei den ca. 150 Fällen von Prostatakarzinomen pro Jahr, die am Schwerpunkt Cytopathologie in Düsseldorf bearbeitet werden, geht es üblicherweise um die Empfehlung, ob man aus DNA-bildzytometrischer Sicht eine gewünschte *Active Surveillance*-Strategie („AS-Strategie“) empfehlen kann oder eher von ihr abraten muss. Dabei wird das Ergebnis der DNA-bildzytometrischen Messung in den Kontext der anderen Kriterien, die gemäß S3-Leitlinie für eine AS-Strategie erfüllt sein sollen (Leitlinienprogramm Onkologie 2016), eingeordnet und als ein „Baustein“ in der Entscheidungsfindung, nicht hingegen als alleiniges Kriterium verstanden. Ein Typ A wird dabei als Indikator für eine Sinnhaftigkeit der AS-Strategie gesehen, ein Typ C und ein Typ D schließen eine Empfehlung dieser Therapieform aus. Beim Typ B wird eine AS-Strategie nicht primär empfohlen, aber als therapeutische Möglichkeit (mit gewisser Erhöhung des Risikos) dargestellt, insbesondere bei älteren Patienten.

1.6 Automatisierte DNA-Bildzytometrie

Üblicherweise werden DNA-bildzytometrische Messungen interaktiv durch einen Untersucher durchgeführt, wobei es sich sowohl um ärztliches Personal, als auch um technisches Personal (MTA, Zytoassistenten) handeln kann. Diese auch als „manuelle Messungen“ bezeichneten Messvorgänge erfolgen somit unter ständiger interaktiver Kontrolle, so dass fehlerhafte Zelldetektionen oder fehlerhafte Zuordnungen von Zellen zu bestimmten Kategorien (Referenzzellen, Analysezellen etc.) von vornherein vermieden werden können. Dieser Vorteil muss allerdings im Kontext des zeitlichen Aufwandes gesehen werden, der abhängig vom Präparat und der Fragestellung beträchtlich sein kann. Außerdem kann stets nur ein Teil der auf dem Objektträger vorhandenen Zellen in die Messung einbezogen werden, so dass (trotz aller Vorgaben in den

Arbeitsanleitungen) stets das Risiko der Ergebnisverfälschung durch Selektion bestimmter visuell auffälliger Zellen besteht.

In einer mehrjährigen Kooperation wurde seit ca. 2006 eine Automatisierung der DNA-Bildzytometrie entwickelt. Kooperationspartner waren dabei der Schwerpunkt Cytopathologie (Wissenschaftliche Leitung: A. Böcking, seit 2010 tätig in Düren), der Lehrstuhl für Bildverarbeitung der RWTH Aachen (Wissenschaftliche Leitung: T. Aach, D. Meyer-Ebrecht) und die Firma Motic (Xianmin, China bzw. Wetzlar). Ziel war, Hardwarekomponenten und eine Software für ein TV-Bildanalysesystem zu schaffen, mit der es möglich werden würde, ein Feulgen-gefärbtes Präparat automatisiert, also ohne primäre Tätigkeit eines Untersuchers, auszuwerten. Die Aufgaben, die dem Fachpersonal zukommen, liegen bei einer solchen Anwendung dann in der Überprüfung der Ergebnisse der Bildanalyse, indem die in einer Bildgalerie abgelegten Objekte auf Plausibilität der Zuordnung zu einzelnen Kategorien (Eichzellen, Analysezellen, normale Zellen, Makrophagen, Artefakte) durchgesehen würden.

Wichtige Fragen, die in diesem Zusammenhang bearbeitet werden mussten, betrafen die Objekterkennung (besonders mit der Thematik der Autofokussierung) und die Objektzuordnung in die genannten Kategorien. Ferner musste das Problem gelöst werden, dass gleiche Zellarten je nach Präparation und eventueller Vorfärbung nicht immer eine identische Morphologie zeigen. So ist beispielsweise ein normaler reifer Lymphozyt in luftgetrockneten Präparationen (z.B. Schilddrüsenpunktaten) wesentlich größer als in mit Alkoholspray fixierten Präparationen (z.B. Bronchialschleimhautausstriche) und noch kleiner in Präparaten, die nach Zellvereinzellung aus Paraffinblöcken von Geweben (z.B. aus Prostatastanzen) gewonnen wurden. Entsprechend musste von Präparation zu Präparation jeweils ein eigener Lerndatensatz generiert werden, anhand dessen das verwendete Computersystem auf eine möglichst hohe Korrelation der Kategorisierung von Zellen im Vergleich zu einem erfahrenen ärztlichen Untersucher gebracht werden konnte. Hierzu wurden von einem Untersucher pro Präparationsart Zehntausende von Objekten detektiert, nach morphologischen Gesichtspunkten kategorisiert und in einer Bilddatenbank abgelegt. Diese Bilddatenbank diente Mathematikern bzw. Informatikern als

Grundlage, um die Software des automatisierten DNA-Bildzytometriesystems immer besser anzupassen. Hierzu wurden komplexe Bildverarbeitungsalgorithmen entwickelt, die später in eine naturwissenschaftliche Promotion eines Mathematikers mündeten (Friedrich 2015). Eine detaillierte wissenschaftliche Erstbeschreibung im medizinischen Kontext wurde im Zusammenhang mit einer oralpathologischen Anwendung veröffentlicht (Böcking et al. 2013).

Das sich hieraus für eine bestimmte Präparation ableitende Klassifikationssystem wird als *Classifier* oder Klassifikator bezeichnet. Derzeit ist nicht nur ein *Classifier* für Ausstriche von Zellsuspensionen nach Zellvereinzelung von Prostatastanzen im Einsatz. Es gibt auch *Classifier* für Alkoholspray-fixierte Abstriche von der Cervix uteri, für Alkoholspray-fixierte Zytozentrifugate von Urin und für luftgetrocknete Sedimentausstriche von Ergüssen (Pleura, Peritoneum, Perikard).

Die Qualität der *Classifier* lässt sich an einem Beispiel aus der Entstehungsphase verdeutlichen (Abb. 7). Für Prostatakarzinome ergab sich bei Berücksichtigung von fast genau 48.000 Objekten eine Übereinstimmung von 90.9% zwischen der Zuordnung, die ein Untersucher vorgenommen hatte, und der Zuordnung durch die entsprechend ausgelegte Software.

	ground truth				
	Cancer Cells	Artifacts	Fibroblasts	Granulocytes	Lymphocytes
classification Cancer Cells	22389	657	192	15	1189
Artifacts	818	15033	257	35	174
Fibroblasts	63	125	1083	1	14
Granulocytes	0	0	0	0	0
Lymphocytes	677	94	8	33	5125
Total	23974	15909	1540	84	6502
Error (%)	6.50	5.51	29.68	100.00	21.18

Abb. 6: Korrelationsmatrix für die Kategorisierung von durch den Untersucher in Gruppen zugeordneten Objekten („ground truth“, Spalten) und der automatisierten Zuordnung („classification“, Zeilen) am Beispiel des *Classifiers* für Prostatakarzinome. Die Übereinstimmung beträgt auf der Basis von 48.009 Objekten 90.9% (Friedrich et al. 2012).

Entsprechend kann man davon ausgehen, dass nunmehr ein *Classifier* vorliegt, der von der sogenannten Papierform her in der Lage sein sollte, eine ausreichend genaue Bestimmung des DNA-zytometrischen Profils eines Prostatakarzinoms zu ermöglichen.

2 Ziel der Arbeit

Was bisher fehlt, ist die klinische Überprüfung der automatisierten Ansätze. Dazu wurden am Schwerpunkt Cytopathologie der Heinrich Heine-Universität Düsseldorf wissenschaftliche Arbeiten initiiert, die eine Überprüfung der bisher etablierten *Classifier* auf Eignung in der täglichen diagnostischen Tätigkeit ermöglichen sollten.

Inzwischen nämlich sind Messeinrichtungen der Firma Motic mit der Option einer automatisierten DNA-Bildzytometrie gewerblich erhältlich, so dass automatisierte Messungen nunmehr auch in die Krankenversorgung eingeführt werden könnten. Am Schwerpunkt Cytopathologie der Heinrich Heine-Universität Düsseldorf wurde eines der ersten automatisierten DNA-Bildzytometriesysteme der Firma Motic etabliert und zur wissenschaftlichen Erprobung freigegeben.

Im Vorfeld klinischer Anwendungen bedarf es methodischer Untersuchungen, um sich über die Messqualität eines solchen Systems im Klaren zu werden und um abzuschätzen, inwieweit die Automatisierung wirklich Vorteile bringt. Insbesondere werden vergleichende Messreihen benötigt, die Ergebnisse manuell durchgeführter Messungen mit denen automatisierter Messungen korrelieren lassen.

In der hier vorliegenden Studie wurde der *Classifier* für Prostata auf seine messtechnischen Qualitäten hin getestet. Hierfür wurden Fälle von Prostatakarzinomen verwendet, an denen bereits manuelle DNA-bildzytometrische Messungen im Rahmen der Routinediagnostik vorgenommen worden waren. Die Messpräparate sollten nunmehr ergänzend automatisierten Messungen unterzogen werden. An diesen sollten vergleichbare Kenndaten für Referenz- und Analysezellen erhoben werden und mit den Ergebnissen der manuellen Messung in Beziehung gesetzt werden.

Ziele dieser Studie lagen darin, unter anderem die technische Qualität des *Classifiers* im Hinblick auf das Ausmaß der richtigen Kategorisierung der Zellen zu testen. Des Weiteren sollte die Dauer und Art der Korrektur von Objekten unter Verwendung der Bildgalerie ermittelt werden. Zum Schluss bestand die Frage, inwieweit sich die Ergebnisse der manuellen und der automatisierten DNA-bildzytometrischen Messungen im Hinblick auf die klinische Aussage korrelieren lassen.

Zu der Studie liegt ein Ethikvotum von der Ethikkommission der Medizinischen Fakultät der Heinrich Heine-Universität Düsseldorf (Studiennummer 4942 vom 30.1.2015) vor.

3 Material und Methoden

3.1 Auswahl der Messpräparate

Wir entschieden uns, die manuelle DNA-Bildzytometrie mit der automatisierten DNA-Bildzytometrie an Fällen von Prostatakarzinomen durchzuführen, die bei der manuellen DNA-Bildzytometrie als Typ B oder als Typ C gradiert worden waren. Diese Vorgabe erfolgte, da genau diese beiden Typen die Grenze zwischen noch eher günstigen Tumoren und möglicher *Active Surveillance-Strategie* bzw. eher ungünstigen Tumoren mit einer konkreten Therapie-notwendigkeit repräsentieren.

Zwischen 2010 und 2015 gab es im Durchschnitt jährlich 65 als Typ B bzw. als Typ C eingeordneter Fälle. Die Zahlen schwankten zwischen 49 und 80 Fällen im Jahr. Im jährlichen Durchschnitt waren 82% von ihnen als Typ B und 18% als Typ C klassifiziert worden.

Insgesamt wurden aus dem Archiv der Cytopathologie am Universitätsklinikum Düsseldorf (UKD) 37 Patienten-Fälle ausgewählt, die im Vorfeld für Fragen der Krankenversorgung manuell gemessen und ausgewertet worden waren und beim Nachmikroskopieren noch technisch einwandfrei erschienen. Ausgewählt wurden die Präparate insbesondere im Hinblick auf die subjektive Stärke der für die DNA-Bildzytometrie notwendigen Feulgen-Färbung, welche technisch bedingt nicht unbegrenzt stabil ist, sondern nach und nach durch natürliches Ausbleichen schwächer wird.

Für die automatisierten Messungen geeignet zeigten sich letztendlich 27 Präparate. Die übrigen 10 waren nicht verwendbar, weil sie entweder zu wenige Eichzellen (Fibroblasten) aufwiesen und / oder doch schon zu stark ausgebleichen waren, so dass sich kein ausreichender Kontrast darstellen ließ.

Zwölf der 27 Fälle wiesen nach manueller Messung in der Routinediagnostik einen Typ B, fünfzehn einen Typ C auf.

Die ersten beiden Schritte wurden dabei bei Raumtemperatur auf einem Schüttler durchgeführt. Durch die Pepsinierung wird ein Verdau von Fasern und Zytoplasmaanteilen erreicht, so dass nunmehr ausschließlich Zellkerne vorlagen. Diese passen durch die Maschenweite der Nylongaze, so dass sie passiv aus dem Nylongazebeutel in die vorgelegte Flüssigkeit übergehen. Vor dem nächsten Schritt wird dann der Beutel vorsichtig entnommen.

4. Herstellung der Präparate: Rörchen 10 min. bei 400g zentrifugieren.
 Überstand bis auf ca. 1 ml abpipettieren.
 Zellsediment auf Poly-L-Lysin beschichtete
 Objektträger auftragen und lufttrocknen.

Die verwendete Nylongaze (Porengröße 50 μm) wurde über die Firma Robert Brückmann (Art.-Nr.: 03-50/31) bezogen, das Pepsin über die Firma Serva (Pepsin Porcine Art.-Nr.: 31820)

3.3 Feulgen-Färbung

Die Feulgen-Färbung (Feulgen und Rossenbeck 1924, Graumann 1953) dient der spezifischen stöchiometrischen Darstellung der DNA. Durch saure Hydrolyse werden die Purinbasen Adenin und Guanin aus den DNA-Molekülen entfernt und hinterlassen freie Aldehydgruppen, an denen sich fuchsinschweflige Säure anlagert. Diese geht eine Bindung mit den freien Aldehydgruppen ein und färbt diese rotlila. Bei RNA hingegen fällt die Reaktion negativ aus, da die Aldehydgruppen räumlich keine Verbindung mit der fuchsinschwefligen Säure eingehen können. Die Färbung kann an zuvor gefärbten Präparaten genauso wie an ungefärbten luftgetrockneten Präparaten angewandt werden. Sie wird automatisiert gemäß den SOPs mit einem Färbeautomaten „Cell Stain“ der Firma Tharmac (Waldsolms, Deutschland) durchgeführt (Tabelle 2, Tabelle 3).

Kritischer Schritt ist die Durchführung der sauren Hydrolyse (Schritt 7, vgl. Tabelle 3), die bei einer definierten Temperatur und unter genauer Einhaltung einer Zeitvorgabe erfolgen muss. Hierzu ist in dem Färbeautomaten eine

heizbare Kuvette eingepasst, deren jeweilige Temperatur einstellbar und kontrollierbar ist. Alle anderen Schritte erfolgen bei Raumtemperatur.

Reagentien	Artikel-Nr.	Hersteller
Ethanol 70%, 96%, 99,5%		UKD-Apotheke
10% gepuffertes Formalin	HT 50-1-128	Sigma
5 N HCl	1.09911.0001	Merck
Schiff'sches Reagenz	1.09033.0500	Merck
1 N HCl	1.09970.0001	Merck
Kaliumdisulfit ($K_2S_2O_5$) zur SO_2 -Wasser-Herstellung	105056	Merck
Xylol	1.08685.2500	Merck
Aqua dest.		UKD-Apotheke

Tab. 2: Geräte und Reagenzien zur Feulgen-Färbung (Arbeitsprotokoll Cytopathologie UKD)

Färbeschritte	Reagenz	Zeit
1.	Xylol	15 min
2.	99,5% Ethanol	5 min
3.	96% Ethanol	5 min
4.	10% gepuffertes Formalin	50 min
5.	Aqua dest.	10 min
6.	Aqua dest.	10 min
7.	5 N HCl (bei 27 °C)	60 min
8.	Aqua dest.	2 min
9.	Aqua dest.	2 min

10.	Aqua dest.	2 min
11.	Schiff's Reagenz	60 min
12.	SO ₂ -Wasser	5 min
13.	SO ₂ -Wasser	5 min
14.	SO ₂ -Wasser	5 min
15.	Aqua dest.	1 min
16.	Aqua dest.	1 min
17.	70% Ethanol	10 min
18.	96% Ethanol	10 min
19.	99,5% Ethanol	10 min
20.	Xylol	15 min

Tab. 3: Ablauf der Feulgen-Färbung (Arbeitsprotokoll Cytopathologie UKD)

Nach Abschluss der Färbung stellt sich ein Zellbild dar, in welchem ausschließlich die Zellkerne der Zellen rotlich angefärbt sind. Das Zytoplasma ist, soweit überhaupt im Rahmen der enzymatischen Verdauung mit Pepsin noch erhalten, ungefärbt bzw. nur schemenhaft erkennbar. Die Präparate werden mit einem konventionellen Eindeckmedium unter einem Deckglas eingedeckt und zum weitmöglichen Schutz gegen Ausbleichen in licht- undurchlässigen Mappen aufbewahrt.

3.4 Manuelle DNA-Bildzytometrie

Die Durchführung der manuellen Messungen erfolgte an einem TV-Bildanalyse-System MotiCyte DNA.

Es besteht aus einem Mikroskop (Motic BA 410 Trinocular, Motic Deutschland, Abb. 8) mit einer 30W-Halogen-Beleuchtung und Plan-Wechselobjektiven; die mikroskopischen Bilder werden mit einer Kamera vom Typ Moticam Pro 285A

(Motic Deutschland, Abb. 8) auf einen Windows PC übertragen und dort mit der Software MotiCyte-Screener analysiert.



Abb. 8: Motic BA410 (Trinocular) 30W Halogen Mikroskop und Moticam Pro 285A Kamera (aus der Bedienungsanleitung des MotiCyte)

Nach Anwählen des Messprogramms wurden die Patientendaten und der zugehörige Patienten-Code eingegeben. Mit der 2x Objektivvergrößerung verschafft man sich danach eine Übersicht über die Zellverteilung auf dem Objektträger und kalibriert das Gerät durch einen Weiß-Abgleich des TV-Kamera-Framegrabber System.

Die eigentliche Messung erfolgt mit der 40x Objektivvergrößerung. Das Präparat wird mäanderförmig abgegangen. Zellkerne, die gemessen werden sollen, werden am Bildschirm mit einer Maus manuell angeklickt und auf diese Weise vom Untersucher ausgewählt. Innerhalb der Analysezellpopulation soll darauf geachtet werden, Zellkerne per Zufall („*random sampling*“) zu erfassen, um eine Verzerrung des Messergebnisses zu vermeiden. Für eine verlässlich interpretierbare Messung werden 300 Analysezellen angestrebt.

Formal wird (vgl. Girmes 2009) beim Messvorgang zuerst die Kontur des zu messenden Objektes auf der Basis des Prinzips der Schwellenwertfindung

detektiert. Danach wird eine Maske um den Kern gelegt und die Kernfläche bestimmt. Danach wird die Integrierte Optische Dichte (IOD) des Messobjektes errechnet, indem innerhalb der Messmaske die Extinktionswerte aller einzelnen Bildpunkte (Pixel) aufsummiert werden. Die Eichung des Systems zur Festlegung des normalen $2c$ -Wertes erfolgt mit internen Referenzzellen, d.h. normalen Zellen, von denen man erwarten darf, dass sie einen diploiden Chromosomensatz besitzen. Die Gleichsetzung der mittleren IOD einer solchen Zellpopulation mit dem normalen $2c$ -Wert ermöglicht, jedem beliebigen IOD-Wert der Analysezellen einem c -Wert zuzuordnen.

Bei der Messung von Prostatakarzinomen werden zur Eichung 30 Fibroblasten herangezogen. Zur technischen Qualitätskontrolle der Eichung muss zum einen darauf geachtet werden, dass der Variationskoeffizient der Messwerte der Referenzzellen möglichst $< 5\%$ liegt und auf diese Weise eine homogene Referenzzellpopulation gegeben ist. Ferner muss der Korrelationskoeffizient zwischen der Kernfläche der Referenzzellen und ihrer IOD $< 0,4$ sein. Beide Größen müssen unkorreliert sein, um störende Streulichteffekte etc. als Verfälschung der Messung ausschließen zu können (vgl. Girmes 2009).

Die Bestimmung der IOD eines Zellkerns erfolgt formal in folgenden Einzelschritten:

- Detektion von Kerngrenzen auf der Basis der Schwellenwertfindung
- Bildung einer Maske um den Kern und Bestimmung der Zellkernfläche
- Erhebung der differentiellen Häufigkeitsverteilung der Grauwerte in der Maske.

Die Differenz aus der mittleren Extinktion des Zellkerns und des Hintergrundes multipliziert mit der Kernfläche ergibt den ungeeichten DNA-Gehalt. Der geeichte DNA-Gehalt errechnet sich aus dem Quotienten des ungeeichten DNA-Gehalts der Analysezelle und dem mittleren Extinktionswert der Eichung.

Aus den Messwerten wird dann vom Programm ein DNA-Histogramm erstellt, das diagnostisch ausgewertet werden kann.

3.5 Automatisierte DNA-Bildzytometrie

Bei der automatisierten DNA-Bildzytometrie (Böcking et al. 2013) werden die Aufgaben, das Präparat zu durchsuchen, die Zellen auszuwählen und zu klassifizieren, vom TV-Bildanalyse-System übernommen. Aufgabe des Untersuchers bleibt, die Auswahl und Klassifikation der Zellen zu kontrollieren und gegebenenfalls zu ändern. Da die Anwendung neuartig ist und am Schwerpunkt Cytopathologie in mehreren wissenschaftlichen Projekten erstmals getestet wird, sollen der technische Ablauf der Messung und die Bearbeitung der Messung durch den Untersucher genauer beschrieben werden.

Das TV-Bildanalyse-System, eine MotiCyte DNA-Workstation (Abb. 9), besteht aus einem motorisierten Mikroskop BA600 (Motic Deutschland GmbH, Wetzlar), gekoppelt mit einer hochauflösenden Moticam Pro 285A 2/3" CCD-Farbkamera (Motic Deutschland GmbH, Wetzlar). Arbeitsobjektiv des Mikroskops ist ein 40x/0,65nA Planobjektiv, wobei die aufgenommenen Bilder 10-fach nachvergrößert werden. Sie werden auf einen Dell Computer Optiplex 9010, der über einen Intel-Core-Prozessor I7-3770 mit 3,4GHz verfügt und unter Windows 7 läuft, übertragen und dort unter Verwendung der Software MotiCyte V2.3 - Screener (Motic Deutschland GmbH, Wetzlar) analysiert.

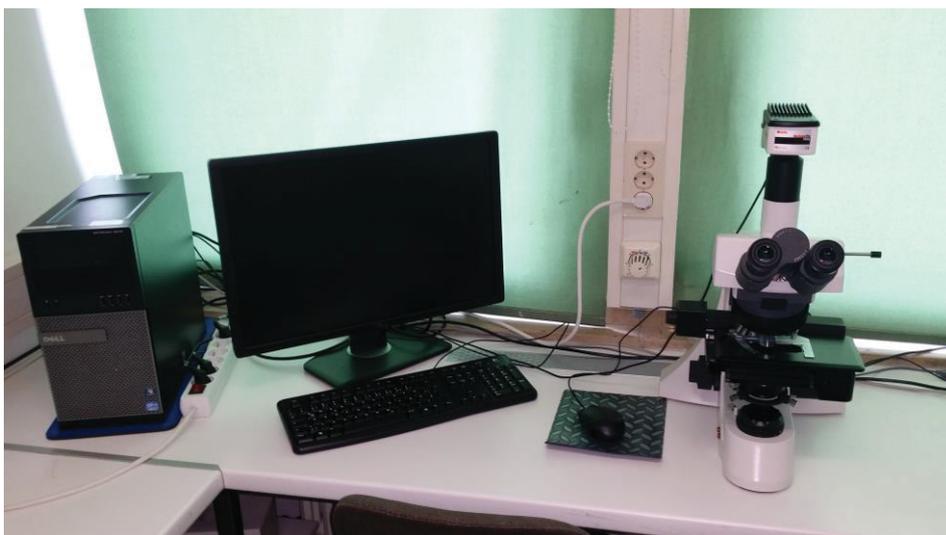


Abb. 9: MotiCyte DNA-Workstation für die automatisierte DNA-Bildzytometrie

Technischer Ablauf der Messung

Wenn der PC hochgefahren ist, wird das Programm durch Doppelklick auf das Symbol "MotiCyte V2.3 - Screener" gestartet. Nach Anlegen einer neuen Patientendatei muss unter dem Reiter "Optionen" und dem Unterpunkt „Programmeinstellungen“ kontrolliert werden, ob der Prostata-Klassifikator ausgewählt ist.

Scannen eines Präparates (Abb. 10, 11 und 12)

Zunächst muss der Beam-Splitter auf die mittlere Position 50% (Okular) / 50% (Kamera) eingestellt werden (Abb. 10). Dann wird die Aperturblende vollständig geöffnet (Abb. 10). Um eine Übersichtskarte zu erzeugen, wird auf den *Button* 'Karte abscannen' geklickt (Abb. 10, Punkt 3). Der Objektivrevolver wird nun auf die 40-fache Vergrößerung eingestellt (Abb. 11, Punkt 4). Die Aperturblende wird soweit geschlossen, dass sie eben gerade nicht mehr im Okular sichtbar ist (Abb. 10). Folgend werden auf dem Präparat Zellen gesucht und diese werden fokussiert. Danach wird das Gerät kalibriert, hierfür wird im Menüreiter der Punkt 'Optionen' und im Unterpunkt 'Kalibration' gewählt (Abb. 11, Punkt 7) und die vorgegebenen Schritte ausgeführt. Nachfolgend wird auf der Übersichtskarte die Region zum Scannen ausgewählt (Abb. 11, Punkt 9).

Im nächsten Schritt werden manuell mindestens 10 sogenannte Fokuspunkte in der für die Messung ausgewählten Region gesetzt. Dazu wird auf der Registerkarte 'Scannen' die Option 'Manuell' ausgewählt (Abb. 12, Punkte 1 und 2). Auf der Übersichtskarte werden 10 verschiedene Punkte ausgewählt und auf die Zellen fokussiert (Abb. 12, Punkt 3). Für jeden Fokus-Punkt muss separat auf die Option 'Fokus-Punkt hinzufügen' angeklickt werden (Abb. 12, Punkt 3). Ziel ist, der Software Anhaltspunkte zu geben, in welchen Bereichen der Z-Achse der Autofokus in bestimmten Regionen des Präparates arbeiten soll.

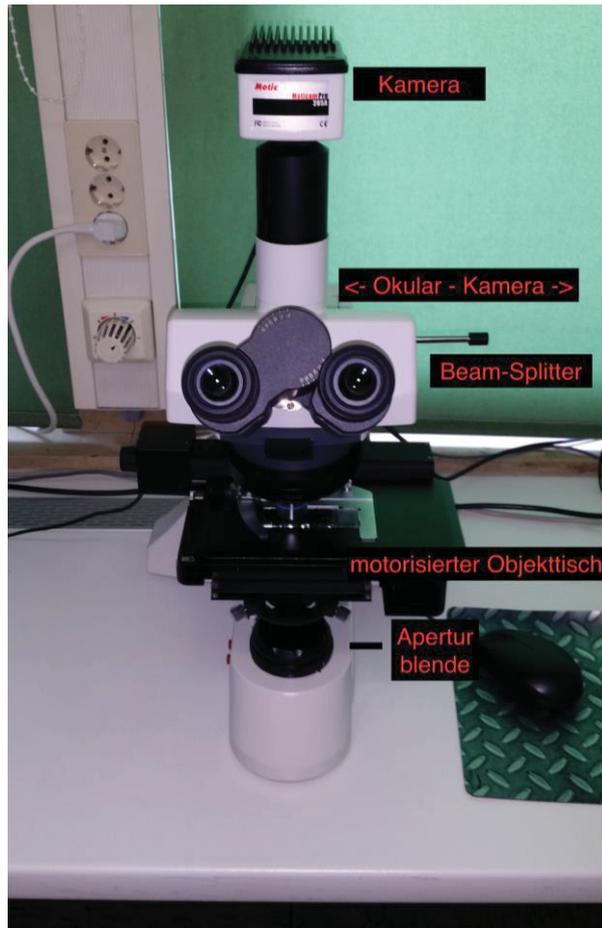


Abb. 10: Automatisiertes Mikroskop (Firma Motic)

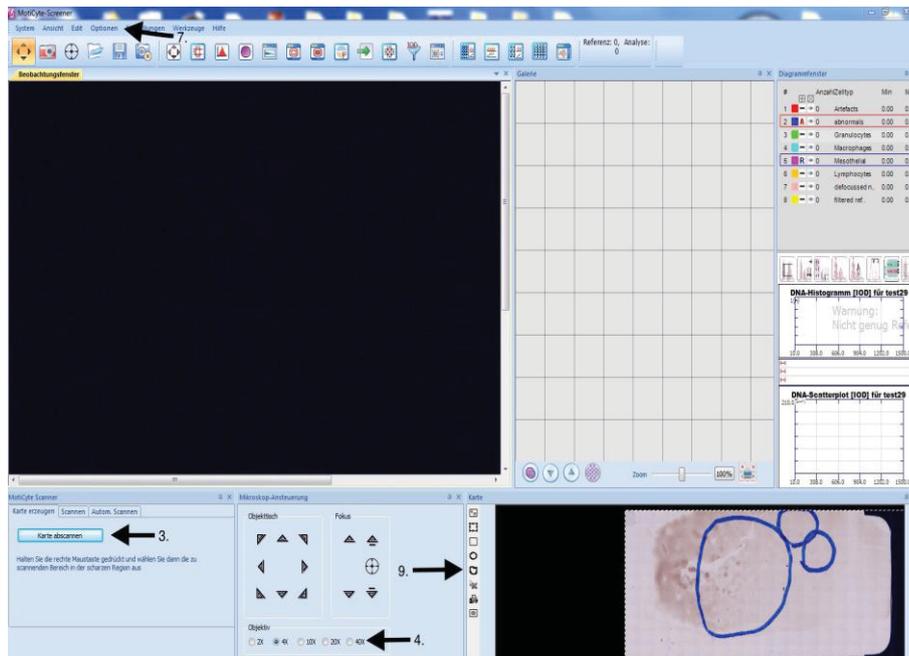


Abb. 11: Screenshot MotiCyte V2.3, Arbeitsschritte Übersichtskarte erzeugen, Kalibrieren und Messregion auswählen.

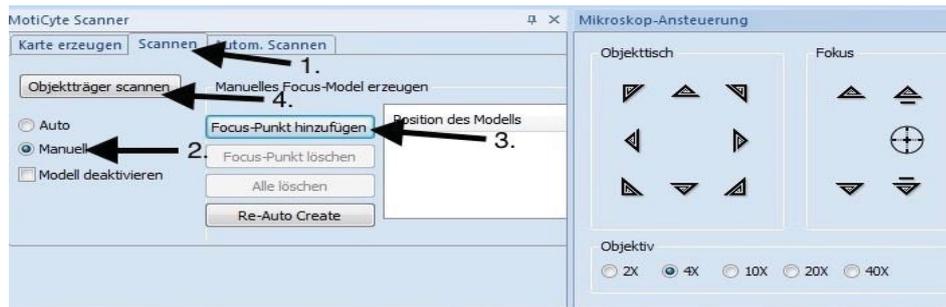


Abb. 12: Screenshot MotiCyte V2.3, Arbeitsschritte Manuelle Fokuspunkte Setzen und Scannen.

Durchführung der automatisierten Messung

Zum Schluss wird das Präparat durch das Anklicken der Option 'Objektträger scannen' vom System automatisch gescannt (Abb. 12, Punkt 4). Das Scannen des Präparats dauert je nach Zellanzahl zwischen 20-45 min. Gemäß den Vorgaben des *Classifiers* werden die erfassten Messobjekte in verschiedenen Unterdateien (Kategorien) der Bildgalerie als Bild abgelegt und als Messwert gespeichert. Es stehen folgende Kategorien zur Verfügung:

Fibroblasts (Referenzzellen)

Zur Eichung dienen normale Fibroblasten, deren Kerne langgestreckt und spitz zulaufend sind (Abb. 14). Ihr DNA-Gehalt ist als 2c anzunehmen.

Abnormals (Analysezellen)

In dieser Kategorie sollen die Tumorzellen abgelegt werden, also diejenigen Messobjekte, die von Form, Größe und Extinktion her als atypische Drüsenzellen der Prostata erkannt werden (Abb. 15). Ihr individueller DNA-Gehalt ist naturgemäß variabel. Untrennbar von abnormen Zellen sind auch in aller Regel normale Epithelien mit einem DNA-Gehalt von 2c untermischt, da gerade bei Prostatakarzinom häufig Tumorareale kleinherdig mit normalem Gewebe

untermischt wachsen. Entsprechend ist in der Analysezellpopulation fast immer ein unterschiedlich großer 2c-Peak vorhanden.

Lymphocytes / Granulocytes (Lymphozyten und Granulozyten)

Lymphozyten und / oder Granulozyten sind den meisten Geweben in gewisser Zahl untermischt und werden bei bestimmten Anwendungen der DNA-Bildzytometrie auch als Referenzzellen verwendet. Bei der automatisierten Messung von Prostatakarzinomen haben sie keine Funktion. Die Zellkerne sind klein und rundlich (Lymphozyten) bzw. klein und gelappt (Granulozyten) und haben einen normalen DNA-Gehalt von 2c.

Artifacts (Artefakte)

Messobjekte, die vom *Classifier* nicht einer bestimmten Zellart zugeordnet werden können, werden als Artefakte eingestuft. Darunter befinden sich Zellkonglomerate und sich überlagernde Zellkerne, aber auch defokussierte Objekte. Auch Analysezellen werden hier gelegentlich irrtümlich abgelegt, z.B. solche mit bizarren Kernformen oder besonders deutlicher Hyperchromasie.

Zwei typische, richtig klassifizierte Messobjekte sind im Folgenden als *Screenshot* aus der Bildgalerie dargestellt (Abb. 13 und Abb. 14).



Abb. 13: Referenzzelle (Fibroblast) mit einem DNA-Gehalt von 2,02c



Abb. 14: Abnorme Zelle mit einem DNA-Gehalt $> 9c$

Bearbeitung der Messung durch den Untersucher

Nach der Messung müssen die Rohdaten bearbeitet werden. Es werden dabei die einzelnen Kategorien in dem Maße auf die Richtigkeit der Zuordnung der Messobjekte untersucht, wie dieses notwendig ist, um sicherzugehen, dass keine messrelevanten Zellen falsch klassifiziert wurden. Eine Durchsicht aller Messobjekte ist hingegen bei der Fülle gemessener Objekte (zum Teil > 20.000) weder möglich, noch sinnvoll. Dazu benötigt der Untersucher morphologische Expertise, um entscheiden zu können, ob ein auffälliges Messobjekt in der richtigen Kategorie eingeordnet ist und in welche Kategorie es möglicherweise stattdessen gehört.

Bei allen Formen geht es aber zunächst darum, eine glaubhafte Eichung zu erreichen, so dass zunächst die Referenzzellen angesehen werden müssen. Bei ihnen ist das Ziel, einen 2c-Peak im Bereich der übrigen normalen Zellen zu ermöglichen, z.B. bei Verwendung von Fibroblasten als Referenzzellen einen 2c-Peak bei Lymphozyten oder innerhalb der Analysezellen einen solchen Peak für die untermischten normalen Epithelien darzustellen. Außerdem muss die Verlässlichkeit der Eichung dadurch gewährleistet sein, dass der Variationskoeffizient (CV) der IOD-Werte für die Referenzzellen bei $< 5\%$ liegen soll. Der Korrelationskoeffizient (CC) zwischen der IOD der Referenzzellen und ihrer Kernfläche soll $< 0,4$ sein; genauer wird auf diese Aspekte in Kapitel 3.7 eingegangen.

Bei der Bearbeitung der Analysezellen ist bei Prostatakarzinomen wesentlich, die DNA-Stammlinien sicher zu identifizieren und dafür zu sorgen, dass alle Messobjekte mit einem Wert von $> 9c$ wirklich einzeln liegende Tumorzellkerne ohne Überlappungen oder sonstige Limitationen sind. Dazu werden bei den Analysezellen die entsprechenden Zellelemente überprüft und ggfls. bei Bedenken in die Artefakte verschoben. Andererseits ist nicht ausgeschlossen, dass sich unter diesen wiederum auch Zellen befinden, die fälschlich als Artefakte klassifiziert worden sind, so dass auch innerhalb dieser Gruppe nach zu berücksichtigenden Zellen gesucht werden muss. Finden sich kleine auffällige Peaks in der Gruppe der Analysezellen, so ist ebenfalls eine Überprüfung der Messobjekte in diesem Bereich vonnöten.

Beispiele für als Artefakte zu klassifizierende Messobjekte zeigen die folgenden *Screenshots* aus der Bildgalerie (Abb. 15 - Abb. 18).

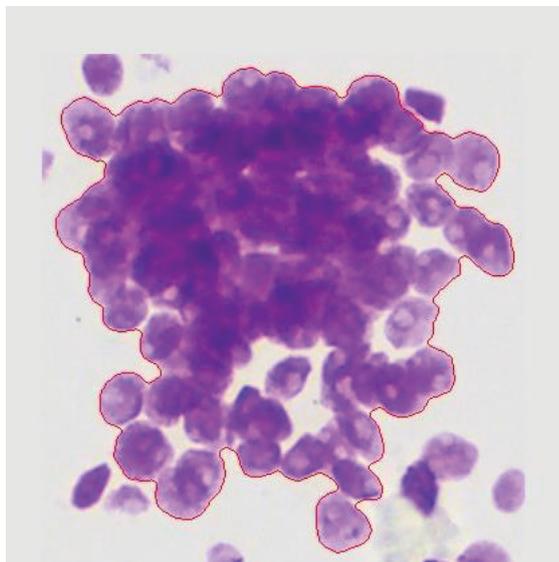


Abb. 15: Zellkonglomerat, das als Artefakt zu klassifizieren ist. Ggfls. könnten bei Bedarf sich nur berührende, aber nicht überlappende Zellkerne (z.B. unten links) über ein Schneidetool getrennt werden und dann der Analysezellgruppe zugeordnet werden.

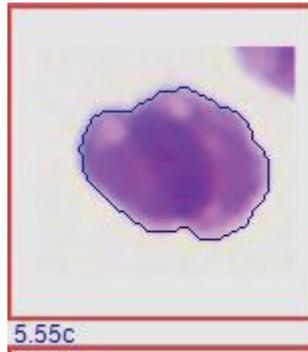


Abb. 16: Dublette aus zwei sich überlappenden Zellkernen, die nicht als Analysezelle interpretiert werden darf, sondern einem Artefakt entspricht.

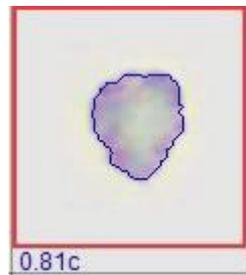


Abb. 17: Nicht interpretierbares, unklares Objekt, das in die Gruppe der Artefakte gehört. Die „ausgefranst“ wirkende Kernkontur und die Aufhellungen im Zentrum sprechen für einen defokussierten Zellkern.

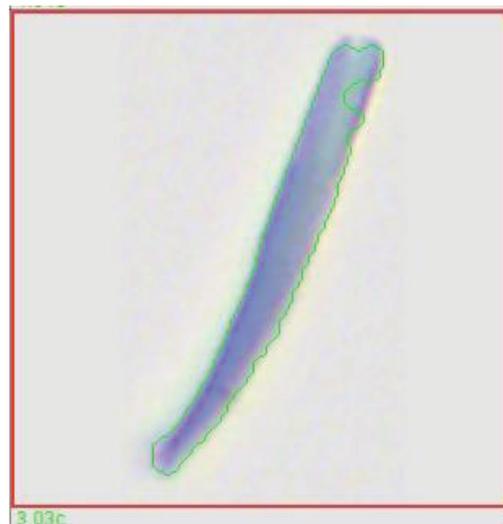


Abb. 18: Ein unklares zelluläres Element, dessen Form zwar an einen Fibroblastenkern erinnert, aber (vgl. Abb. 13 in gleicher Vergrößerung) viel zu groß hierfür ist. Auch dieses Element kann nur als Artefakt klassifiziert werden.

Auswertung der Messdatei nach Bearbeitung der Bildgalerie

An die Bearbeitung der Bildgalerie schließt sich die Extraktion der Daten für die Erstellung des DNA-Histogramms und für statistische Zwecke an. Die einzelnen Parameter, die in der hier vorliegenden Arbeit verwendet wurden, sind dabei in Kapitel 3.7 dargelegt. Im Folgenden sind zwei Histogramme, jeweils eines für einen Typ B nach Tribukait bzw. für einen Typ C nach Tribukait, wiedergegeben (Abb. 19 und Abb. 20).

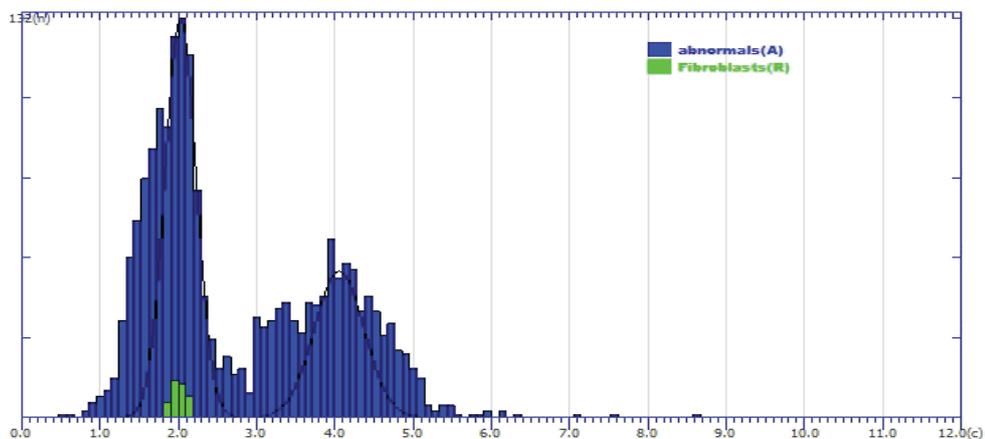


Abb. 19: Histogramm Typ B mit Stammlinien bei 2,02c und 4,0c; keine Zellen mit einem DNA-Gehalt > 9c

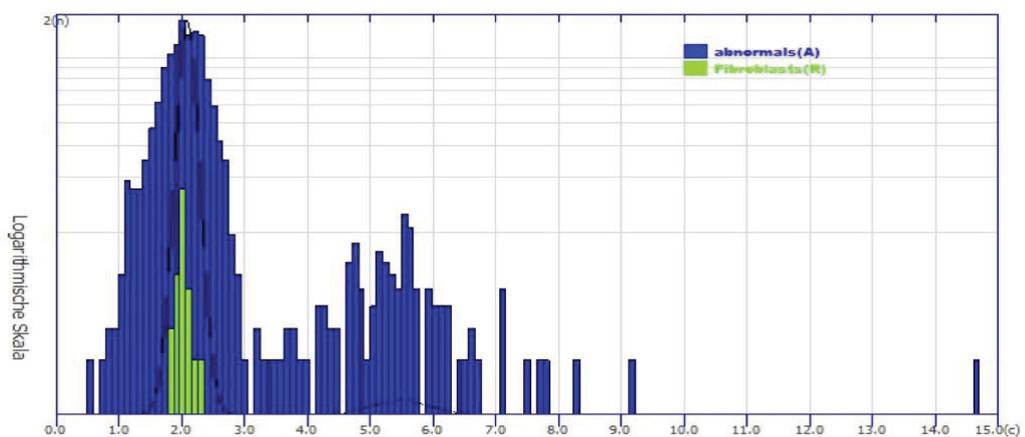


Abb. 20: Histogramm Typ C nach Tribukait. Zwei DNA-Stammlinien bei 2,1c, 5,5c sowie eine Zelle mit einem DNA-Gehalt > 9c

3.6 Aufbau der Messreihe

Jeder Patientenfall wurde dreimal gemessen.

- Aus dem Fallarchiv des Schwerpunktes Cytopathologie stand das Ergebnis der manuell vorgenommenen konventionellen DNA-bildzytometrischen Messung aus der Krankenversorgung zur Verfügung.
- Bei der ersten automatisierten Messung wurde das System hochgefahren und der automatisierte Messvorgang, wie in den obigen Schritten erklärt, durchgeführt.
- Bei der zweiten automatisierten Messung wurde das komplette System wieder heruntergefahren, so dass alle Voreinstellungen der vorherigen Messung gelöscht waren, danach neu gestartet und eine vollkommen neue Messung angeschlossen.

3.7 Untersuchte Messparameter

Im Folgenden sollen die untersuchten Messparameter charakterisiert werden. Dabei müsste streng genommen stets zwischen „Objekten“ in den Rohdaten und „Zellen“ bzw. „Zellkernen“ in den kontrollierten Daten unterschieden werden, was aus Gründen der Praktikabilität hier nicht durchgehend verwirklicht werden konnte.

Zahl der Referenzzellen und der Analysezellen

Die Zahlenangaben zu den Referenzzellen und den Analysezellen vor und nach der Kontrolle der Bildgalerie wurden in allen Fällen dokumentiert. Nicht berücksichtigt wurde, wie viele Objekte eliminiert bzw. hinzugewonnen wurden. Es kann somit nicht unterschieden werden, ob eine Veränderung von z.B. 540 Zellen auf 536 Zellen bedeutet, dass genau vier Objekte aus dieser Klasse in eine andere verschoben worden sind. Genauso gut könnten zehn Objekte nicht akzeptiert worden sein und sechs andere Objekte durch eine Zuweisung aus einer anderen Klasse von Zellen hinzugekommen sein.

Variationskoeffizient (CV) der integrierten optischen Dichte (IOD) der Referenzzellen

Um von einer verlässlichen Eichung ausgehen zu können, ist üblich, bei der IOD der Eichzellen einen Variationskoeffizienten von $< 5\%$ zu fordern. Ist dieses nicht realisierbar, so kann man das Messergebnis nur bedingt als valide ansehen. Ob es akzeptiert wird, bleibt aber in der klinischen Routine dem Untersucher überlassen. Bei der Kontrolle der Bildgalerie wurde daher genau darauf geachtet, Fehlzuordnungen bei den Referenzzellen aufzufinden und zu korrigieren.

Korrelationskoeffizient (CC) zwischen der IOD und der Fläche der Referenzzellen

Auf DNA-bildzytometrische Messergebnisse können auch physikalische Einflussgrößen entscheidend einwirken, zu denen Effekte wie *Blooming*, *Glare*, *Shading* oder Refraktionsphänomene gehören. Im weitesten Sinne können die hieraus resultierenden Probleme heute über Software-Anwendungen gelöst werden. Ein wichtiger Kontrollparameter dafür, ob man die Optik im Griff hat, ist der Korrelationskoeffizient zwischen der IOD und der Fläche der Referenzzellen, im Folgenden mit CC abgekürzt.

Ziel ist nachzuweisen, dass IOD-Werte der oft kleinen oder schmalen Referenzzellen unabhängig von ihrer Kernfläche sind. Stellt man sich dieses graphisch vor, indem man auf der x-Achse die Kernfläche und auf der y-Achse die IOD aufträgt, so sollen die Messwertpaare in einem parallel zu den Achsen liegenden Rechteck dargestellt sein. Ab einer bestimmten Kleinheit der Referenzzellkerne beobachtet man aber, dass eine Abhängigkeit beider Größen insoweit einsetzt, dass ein Richtung (0/0) gerichteter Shift einsetzt; an das besagte Rechteck schließt sich dann eine Art Schwanz zum Nullpunkt des Koordinatensystems an, in welchem eine am Ende lineare Abhängigkeit zwischen Kernfläche und IOD besteht. Liegt der CC-Wert zwischen beiden Parametern $< 0,4$, so geht man von einer Unabhängigkeit der Parameter und damit einer verlässlichen Eichung aus.

Inhaltlich liegt das Problem darin, dass bei zu kleinen Eichzellen die IOD-Werte durch Beugung von Licht in den Strahlengang an der Kernkontur zu niedrig ausfallen, somit der 2c-Wert für die Referenzzellen auf einem zu niedrigen mittleren IOD-Wert basiert und konsekutiv eine Rechtsverschiebung der DNA-Werte der Analysezellen resultiert. Hierdurch kann es zu falsch aneuploiden DNA-Stammlinienbestimmungen und somit zur klinischen Fehlinterpretation der Messung kommen (vgl. Girmes 2009).

Abbildung 21 zeigt einen Ausdruck dieser methodischen Untersuchung, wie er Teil der Dokumentation einer manuellen DNA-bildzytometrischen Messung ist.

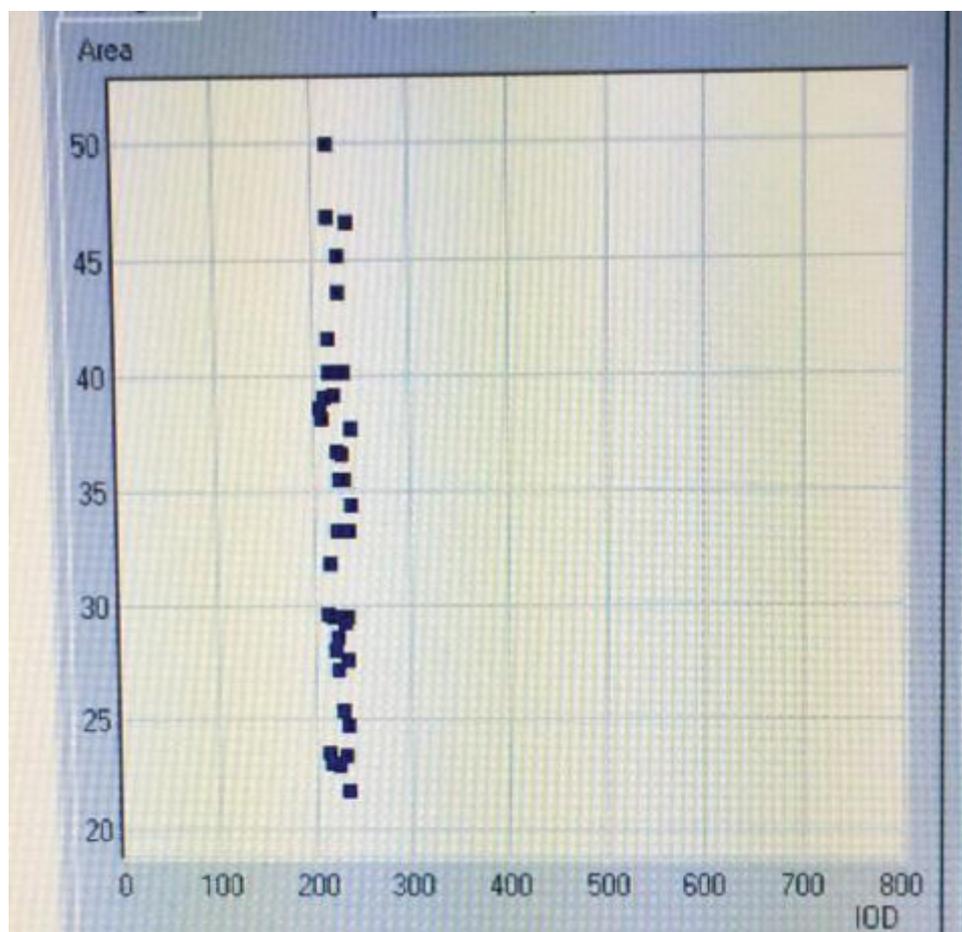


Abb. 21: Dokumentation zur Qualität der Eichung bei einer manuellen DNA-bildzytometrischen Messung. Aufgetragen sind auf der x-Achse die IOD-Werte der Eichzellen und auf der y-Achse die Kernfläche der Eichzellen in μm^2 . Die resultierende Punktwolke liegt in einem imaginären Rechteck ohne Hinweis auf eine lineare Beziehung zwischen beiden Variablen. CC-Wert: 0,02 (Biesterfeld 2014).

DNA-Stammlinie

Die DNA-Stammlinie, als Begriff 1963 eingeführt von Sandritter (Seidel und Sandritter 1963, Sandritter et al. 1966), benennt den DNA-Gehalt der Zellen der G₀/G₁-Fraktion einer Zellpopulation und müsste im Idealfall normaler Zellen, z.B. Fibroblasten, Lymphozyten oder regelrechter Prostataepithelien, bei oder sehr nahe bei 2,0c liegen.

Je nach Arbeitsgruppe liegen leicht unterschiedliche Definitionen vor, insbesondere abhängig davon, ob sich die Lage der Stammlinie auf den Median, den Mittelwert oder den Modalwert der DNA-Verteilung im zu berücksichtigenden Bereich der Messwerte bezieht. In dieser Studie bestimmte sich die DNA-Stammlinie als der Modalwert der DNA-Verteilung der gemessenen Zellkerne. Sie wurde als diploid definiert, wenn der Modalwert innerhalb des 97,5%-Konfidenzintervalls des mittleren DNA-Gehaltes der Eichzellpopulation lag (= „diploide DNA-Stammlinie“), als peritetraploid, wenn der Modalwert sich innerhalb des doppelten Vielfachen des 95%-Konfidenzintervalls des mittleren DNA-Gehaltes der Eichzellpopulation befand (= „peritetraploide DNA-Stammlinie“) und als aneuploid in denjenigen Fällen, die einen Modalwert außerhalb der diploiden und peritetraploiden Bereiche zeigten (= „aneuploide DNA-Stammlinie“).

Exceeding Events

Werte für die „*Exceeding Events*“ (EE) geben an, wie viele Zellen einen bestimmten absoluten DNA-Wert, z.B. 5c oder 9c, überschreiten und werden als z.B. 5cEE bzw. 9cEE bezeichnet. Hinter diesen beiden Variablen steht die biologische Überlegung, DNA-Messwerte zu quantifizieren, die die Grenzen für den Normalwert der G₂/M-Phase (4c) oder der G₂/M'-Phase (8c) eindeutig überschreiten. In dieser Studie wurde der 9cEE berücksichtigt; Messungen wurden u.a. als Tribukait C kategorisiert, wenn sie einen 9cEE > 0 aufwiesen.

C-Max

Der C-Max stellt den maximalen DNA-Einzelwert einer Analysezelle in der gesamten Analysezellpopulation dar.

3.8 Statistik

Die statistischen Untersuchungen wurden in Kooperation mit dem Institut für Pathologie in Koblenz unter Verwendung des biometrischen Programmpaketes BMDP und seiner im Folgenden in Klammern genannten Unterprogramme vorgenommen. Dazu gehörten die deskriptive Grundstatistik mit Errechnung von Mittelwerten, Medianen und Datenverteilungen etc. (1D, 2D), ferner bei kategorisierten Daten die Berechnung von Kreuztabellen (Mehr-Felder-Tafeln) und ihre Auswertung mit dem Pearson Chi²-Test (4F) bzw. mit Cohen's Kappa sowie bei nicht gestuften Daten die Berechnung von Korrelationskoeffizienten und die Durchführung eines t- bzw. U-Tests (3D). Streudiagramme wurden zur graphischen Veranschaulichung genutzt.

Zur Darstellung der Tendenz und Größe der Abweichung von Messergebnissen erfolgte bei den Untersuchungen zu den Peaks außerdem eine erweiterte statistische Analyse und eine graphische Darstellung mit Bland-Altman-Diagrammen am Schwerpunkt Cytopathologie des Universitätsklinikums Düsseldorf.

Mit der Hilfe des Kolmogorov-Smirnov-Tests und des Shapiro-Wilk-Tests wurde hier zunächst auf Normalverteilung der Daten überprüft. Sofern eine Normalverteilung angenommen werden konnte, wurde der Spearman'sche Korrelationskoeffizient berechnet, und zwar zur Überprüfung, ob eine Korrelation zwischen den Abweichungen der Messwerte untereinander (*Bias*) und den Mittelwerten der Messwerte besteht. Als H_0 (Nullhypothese) wurde angenommen, dass keine signifikante Korrelation zwischen dem *Bias* und dem Mittelwert besteht. Außerdem wurde der Student t-Test angewendet, um die Nullhypothese zu testen, dass der *Bias* gleich Null ist.

Abschließend wurden die Daten auf ein Bland-Altman-Diagramm übertragen. Hierbei werden pro Messwertpaar der Durchschnitt der beiden Messwerte auf der x-Achse und die Differenz der beiden Messwerte auf der y-Achse aufgetragen. Abb. 22 zeigt hierfür ein Beispiel.

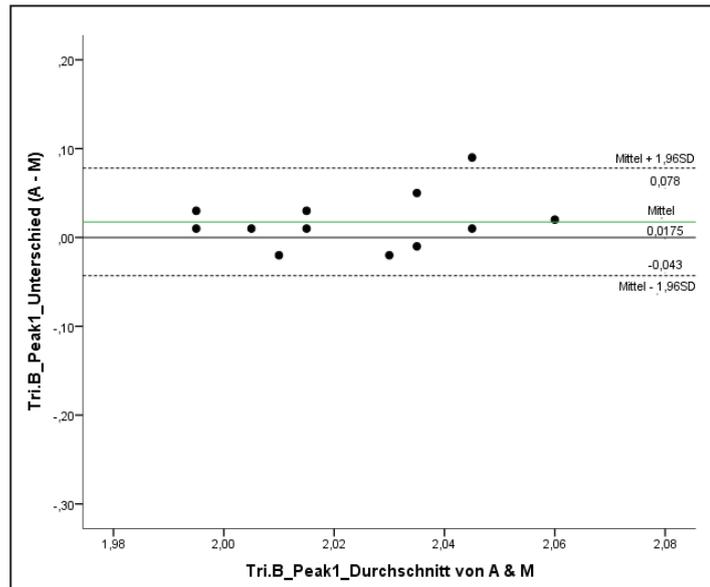


Abb. 22: Beispiel für ein Bland-Altman-Diagramm für 12 Messwertpaare aus zwei Messreihen. Aufgetragen sind auf der x-Achse die Mittelwerte aus den jeweiligen beiden Messwerten und auf der y-Achse ihre Differenz. Eingetragene Hilfslinien betreffen die ideale Linie bei übereinstimmenden Messwerten entsprechend einer Differenz von Null (_____), die Linie für den Mittelwert (*Bias*) der wirklich ermittelten Differenzen (_____) und Linien für die obere bzw. untere Grenze des 95%-Konfidenzintervalls der Differenzen (- - - - -).

Sofern Testungen angewandt wurden, galt ein p-Wert von $p \leq 0,05$ als statistisch signifikant, entsprechend einer Irrtumswahrscheinlichkeit von $\leq 5\%$, die Nullhypothese fälschlich abzulehnen.

Die oben angesprochene Anwendung der Kappa-Statistik führt zu Werten zwischen 0,0 und 1,0. κ -Werte $< 0,20$ zeigen eine schwache Übereinstimmung an, κ -Werte zwischen 0,21 und 0,4 eine leichte, zwischen 0,41 und 0,6 eine mäßige, zwischen 0,61 und 0,8 eine gute und von $> 0,8$ an eine sehr gute Übereinstimmung.

4 Ergebnisse

4.1 Patientenkollektiv

Von den 37 Messpräparaten konnten 27 statistisch ausgewertet werden, darunter 12 Tribukait B- und 15 Tribukait C-Fälle. Die übrigen 10 Präparate waren bereits zu stark abgeblasst bzw. enthielten zu wenige Eichzellen, um noch für weitere Untersuchungen zur Verfügung zu stehen. Versuche, sie zu messen, führten zu keinem methodisch unanfechtbaren Resultat. In einem Fall konnte nur die erste der beiden vorgesehenen automatisierten Messungen erfolgen, so dass er in den methodischen Untersuchungen nicht berücksichtigt wurde; diese beziehen sich somit auf 52 Messungen von 26 Fällen.

4.2 Methodische Untersuchungen an den Referenzzellen

4.2.1 Zahl der Referenzzellen

In Tabelle 4 sind für die Zahl der als Referenzzellen erkannten Messobjekte die Rohdaten wiedergegeben. Tabelle 5 zeigt die statistischen Kenndaten.

Fall	Messreihe	Referenzzellen vor Bearbeitung der Bildgalerie	Referenzzellen nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
1	1	23	36	-13.00	-56.52
1	2	251	40	211.00	84.06
2	1	324	57	267.00	82.41
2	2	332	37	295.00	88.86
3	1	526	85	441.00	83.84
3	2	478	74	404.00	84.52
4	1	124	39	85.00	68.55
4	2	122	92	30.00	24.59
5	1	414	42	372.00	89.86
5	2	378	62	316.00	83.60
6	1	97	72	25.00	25.77
6	2	94	30	64.00	68.09

7	1	65	33	32.00	49.23
7	2	85	31	54.00	63.53
8	1	481	41	440.00	91.48
8	2	454	51	403.00	88.77
9	1	303	71	232.00	76.57
9	2	518	80	438.00	84.56
10	1	136	70	66.00	48.53
10	2	126	80	46.00	36.51
11	1	106	45	61.00	57.55
11	2	105	42	63.00	60.00
12	1	22	38	-16.00	-72.73
12	2	22	30	-8.00	-36.36
13	1	101	33	68.00	67.33
13	2	91	32	59.00	64.84
14	1	82	19	63.00	76.83
14	2	85	30	55.00	64.71
15	1	35	35	0.00	0.00
15	2	43	30	13.00	30.23
16	1	249	133	116.00	46.59
16	2	152	74	78.00	51.32
17	1	111	35	76.00	68.47
17	2	41	35	6.00	14.63
18	1	244	30	214.00	87.70
18	2	187	33	154.00	82.35
19	1	627	157	470.00	74.96
19	2	625	165	460.00	73.60
20	1	31	30	1.00	3.23
20	2	28	30	-2.00	-7.14
21	1	111	30	81.00	72.97
21	2	98	35	63.00	64.29
22	1	20	30	-10.00	-50.00
22	2	24	32	-8.00	-33.33
23	1	157	30	127.00	80.89
23	2	191	43	148.00	77.49
24	1	71	30	41.00	57.75
24	2	64	30	34.00	53.13
25	1	244	78	166.00	68.03
25	2	249	57	192.00	77.11
26	1	34	66	-32.00	-94.12
26	2	298	137	161.00	54.03

Tab. 4: Datenübersicht über die Referenzzellen vor und nach Bearbeitung der Bildgalerie für beide Messreihen à 26 Messungen (n = 52)

Messreihe	Referenzzellen vor Bearbeitung der Bildgalerie	Referenzzellen nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
Beide Messreihen				
Mittelwert	189,98	53,40	136,58	48,14
SD	168,95	33,32	150,00	45,67
CV	88,93	62,40	109,83	94,87
Median	116,50	38,50	67	64,77
Kleinster Wert	20	19	-32	-94,12
Größter Wert	627	165	470	91,48
Messreihe 1				
Mittelwert	182,23	52,50	129,73	42,51
SD	169,91	32,76	152,18	54,00
CV	93,24	62,41	117,30	127,03
Median	111	38,50	72	67,68
Kleinster Wert	20	19	-32	-94,12
Größter Wert	627	157	470	91,48
Messreihe 2				
Mittelwert	197,73	54,31	143,42	53,77
SD	171,00	34,50	150,48	35,68
CV	86,47	63,52	104,92	66,36
Median	124	38,50	63,50	64,49
Kleinster Wert	22	30	-8	-36,36
Größter Wert	625	165	460	88,86

Tab. 5: Statistische Grunddaten über die Referenzzellen vor und nach Bearbeitung der Bildgalerie. Angaben für beide Messreihen à 26 Messungen (n = 52) sowie für die einzelnen Messreihen 1 und 2 getrennt. SD = Standardabweichung, CV = Variationskoeffizient

Rein deskriptiv lässt sich aus den Daten ableiten, dass in aller Regel die Zahl der als Referenzzellen berücksichtigten Objekte bei der Durchsicht der Bildgalerie vermindert wurde, was der Erwartung entspricht. Nur bei vier der 52 Messungen (7,7%) wurde die Zahl der Eichzellen durch Hinzunahme von Objekten erhöht. Die hohen Werte für SD und CV verdeutlichen, dass die Bearbeitung der Bildgalerie individuell erfolgte, auch wenn die Vorgaben, nach denen Referenzzellen ausgewählt werden sollten, stets die gleichen waren.

In Tabelle 6 sind die Korrelationskoeffizienten zwischen den einzelnen Variablen dokumentiert. Mit Werten von 0,6095 bzw. 0,6506 sind die Zahlen der Referenzzellen vor und nach Bearbeitung der Bildgalerie noch signifikant korreliert. Sehr hoch korreliert waren die Zahlen der Referenzzellen vor der Bearbeitung und die Differenz: Je mehr Objekte vorhanden waren, desto mehr Zellen wurden in der Regel auch durch den Untersucher eliminiert.

Referenzzellen	vor Bearbeitung	nach Bearbeitung	Differenz	Differenz in %
Messreihe 1				
vor Bearbeitung	"---			
nach Bearbeitung	0.6095	"---		
Differenz	0.9853	0.4653	"---	
Differenz in %	0.5925	0.1090	0.6381	"---
Messreihe 2				
vor Bearbeitung	"---			
nach Bearbeitung	0.6506	"---		
Differenz	0.9844	0.5068	"---	
Differenz in %	0.6260	0.1376	0.6787	"---

Tab. 6: Korrelationskoeffizienten für die zur Zahl der Referenzzellen erhobenen Variablen, getrennt nach den Messreihen 1 und 2.

Sowohl im t-Test, als auch im Mann-Whitney U-Test ergab sich im Vergleich beider Messreihen keine statistische Signifikanz ($p > 0,05$) (Tabelle 7).

Referenzzellen	t-Test	Mann-Whitney U-Test
vor Bearbeitung	0.6388	0.5279
nach Bearbeitung	0.7703	0.7189
Differenz	0.6432	0.6785
Differenz in %	0.2369	0.5719

Tab. 7: p-Werte des t-Tests und des Mann-Whitney U-Tests im Vergleich der Messreihen 1 und 2 für die zur Zahl der Referenzzellen erhobenen Variablen

Für die 52 Messungen zeigt Abb. 23 die Beziehung zwischen der Zahl der Referenzzellen vor und nach Bearbeitung der Bildgalerie. Auch wenn sich keine ideale Gerade darstellt, zeigt sich doch im Wesentlichen eine Linearität der Ergebnisse, was für eine vergleichbare Vorgehensweise bei der Bildbearbeitung spricht.

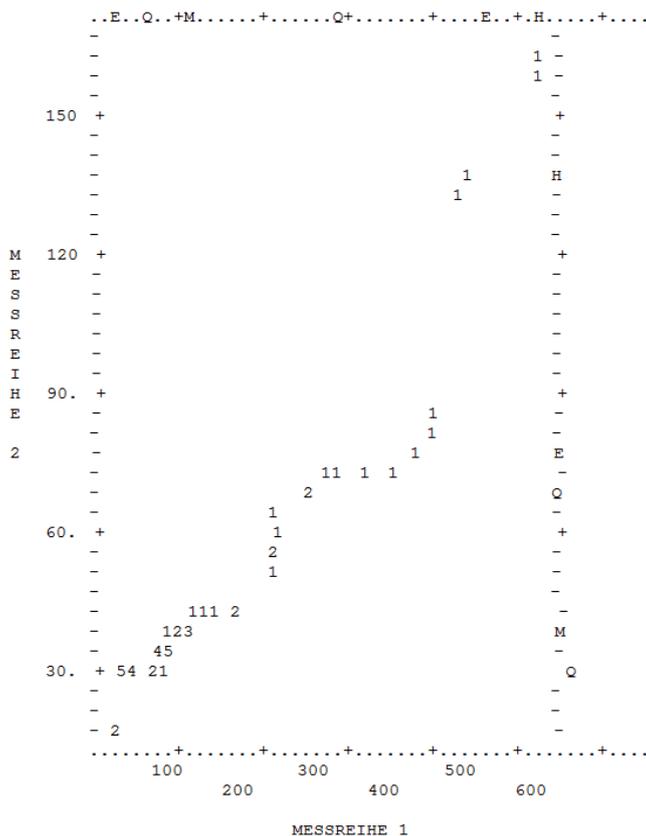


Abb. 23: Graphische Darstellung der 52 Messwerte der Messreihen 1 und 2 für die Zahl der Referenzzellen vor und nach Bearbeitung der Bildgalerie. Die Ziffern 1 bis 5 in der Grafik geben wieder, wie viele Messpunkte auf einen Datenpunkt fallen. Die Ziffernsumme entspricht somit 52.

4.2.2 Variationskoeffizient (CV) der integrierten optischen Dichte (IOD) der Referenzzellen

In Tabelle 8 sind die Variationskoeffizienten für die IOD der Referenzzellen fallweise dargestellt, in Tabelle 9 die zugehörigen statistischen Kenndaten.

Fall	Messreihe	CV-Wert vor Bearbeitung der Bildgalerie	CV-Wert nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
1	1	31.10	4.95	26.15	84.08
1	2	27.88	4.47	23.41	83.97
2	1	15.06	4.58	10.48	69.59
2	2	14.42	3.63	10.79	74.83
3	1	17.65	2.79	14.86	84.19
3	2	18.60	2.42	16.18	86.99
4	1	17.62	3.77	13.85	78.60
4	2	18.30	4.00	14.30	78.14
5	1	18.59	1.34	17.25	92.79
5	2	20.67	3.13	17.54	84.86
6	1	10.39	4.02	6.37	61.31
6	2	14.84	4.98	9.86	66.44
7	1	38.30	3.10	35.20	91.91
7	2	28.49	3.94	24.55	86.17
8	1	18.99	2.82	16.17	85.15
8	2	18.18	3.24	14.94	82.18
9	1	8.65	1.15	7.50	86.71
9	2	16.73	1.36	15.37	91.87
10	1	15.04	2.85	12.19	81.05
10	2	13.38	2.75	10.63	79.45
11	1	30.41	2.92	27.49	90.40
11	2	10.81	1.66	9.15	84.64
12	1	22.42	4.91	17.51	78.10
12	2	16.86	4.96	11.90	70.58
13	1	20.38	4.92	15.46	75.86
13	2	17.58	4.97	12.61	71.73
14	1	20.75	4.69	16.06	77.40
14	2	24.13	11.97	12.16	50.39
15	1	12.52	4.74	7.78	62.14
15	2	15.03	3.86	11.17	74.32
16	1	14.76	4.00	10.76	72.90
16	2	14.58	4.07	10.51	72.09
17	1	21.88	4.99	16.89	77.19
17	2	8.72	4.98	3.74	42.89
18	1	19.95	4.94	15.01	75.24
18	2	22.79	4.78	18.01	79.03
19	1	11.39	2.67	8.72	76.56
19	2	11.97	3.25	8.72	72.85
20	1	34.82	21.55	13.27	38.11
20	2	31.15	19.04	12.11	38.88
21	1	28.74	4.99	23.75	82.64
21	2	29.72	4.69	25.03	84.22
22	1	20.16	4.89	15.27	75.74
22	2	11.68	3.87	7.81	66.87

23	1	24.07	4.89	19.18	79.68
23	2	16.02	3.92	12.10	75.53
24	1	22.48	4.67	17.81	79.23
24	2	27.24	6.00	21.24	77.97
25	1	12.03	3.50	8.53	70.91
25	2	14.87	2.15	12.72	85.54
26	1	28.14	3.61	24.53	87.17
26	2	14.90	3.66	11.24	75.44

Tab. 8: Datenübersicht über die CV-Werte vor und nach Bearbeitung der Bildgalerie für beide Messreihen à 26 Messungen (n = 52)

Messreihe	CV-Wert vor Bearbeitung der Bildgalerie	CV-Wert nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
Beide Messreihen				
Mittelwert	19,54	4,62	14,92	76,01
SD	6,99	3,55	6,17	12,15
CV	35,77	76,85	41,36	15,99
Median	18,24	4,00	14,08	78,04
Kleinster Wert	8,65	1,15	3,74	38,11
Größter Wert	38,30	21,55	35,20	92,79
Messreihe 1				
Mittelwert	20,63	4,55	16,08	77,49
SD	7,63	3,65	6,88	11,26
CV	37,01	80,20	42,79	14,53
Median	20,06	4,30	15,37	78,35
Kleinster Wert	8,65	1,15	6,37	38,11
Größter Wert	38,30	21,55	35,20	92,79
Messreihe 2				
Mittelwert	18,44	4,68	13,76	74,53
SD	6,23	3,51	5,25	13,03
CV	33,80	75,06	38,14	17,48
Median	16,80	3,93	12,14	76,75
Kleinster Wert	8,72	1,36	3,74	38,88
Größter Wert	31,15	19,04	25,03	91,87

Tab. 9: Statistische Grunddaten der CV-Werte vor und nach Bearbeitung der Bildgalerie. Angaben für beide Messreihen à 26 Messungen (n = 52) sowie für die einzelnen Messreihen 1 und 2 getrennt. SD = Standardabweichung, CV = Variationskoeffizient

Man erkennt, dass das Ziel der Bearbeitung durch den Untersucher, nämlich die Senkung des CV-Wertes auf $< 5\%$, bei 48 der 52 Messungen (92,3%), erreicht werden konnte. Bei vier Fällen gelang dieses aber trotz aller Bemühungen nicht. Die Reduktion der Zellzahl um im Durchschnitt jeweils knapp $> 75\%$ belegt, dass die automatisiert als Eichzellen vorgesehenen Messobjekte doch eine gewisse Messwertschwankung zeigen, die eine Beschränkung der Eichzellen auf eine vergleichsweise kleine Stichprobe notwendig erscheinen ließ.

In Tabelle 10 sind die Korrelationskoeffizienten der untersuchten Variablen angegeben. Die eher geringe Korrelation zwischen den Referenzzellen vor und nach Bearbeitung der Bildgalerie (0,4353 bzw. 0,5416) belegt, dass von Fall zu Fall individuell vorgegangen werden musste und ein immer gleiches Vorgehen nicht anzunehmen ist.

CV-Wert	vor Bearbeitung	nach Bearbeitung	Differenz	Differenz in %
	Messreihe 1			
vor Bearbeitung	"---			
nach Bearbeitung	0.4353	"---		
Differenz	0.8788	-0.0471	"---	
Differenz in %	0.1158	-0.7913	0.5480	"---
	Messreihe 2			
vor Bearbeitung	"---			
nach Bearbeitung	0.5416	"---		
Differenz	0.8256	-0.0272	"---	
Differenz in %	0.0153	-0.7709	0.5358	"---

Tab. 10: Korrelationskoeffizienten für die zum Variationskoeffizienten der Referenzzellmessung (CV) erhobenen Variablen, getrennt nach den Messreihen 1 und 2.

Die Ergebnisse des t-Tests und des Mann-Whitney U-Tests für den Vergleich der Messreihen 1 und 2 zeigen für alle vier untersuchten Variablen keine

Signifikanz ($p > 0,05$), so dass sich keine Anhaltspunkte dafür ergeben, dass die Daten nicht zu einer einheitlichen Stichprobe gehören (Tabelle 11).

CV	t-Test	Mann-Whitney U-Test
vor Bearbeitung	0.2875	0.2504
nach Bearbeitung	0.9035	0.8211
Differenz	0.2008	0.2543
Differenz in %	0.4206	0.4740

Tab. 11: p-Werte des t-Tests und des Mann-Whitney U-Tests im Vergleich der Messreihen 1 und 2 für die zum Variationskoeffizienten (CV) der Referenzzellen erhobenen Variablen

4.2.3 Korrelationskoeffizient (CC) zwischen der mittleren IOD und der Kernfläche der Referenzzellen

Die Datenverteilung der CC-Werte und die statistischen Angaben zu Mittelwerten, Medianen etc. sind in den Tabellen 12 und 13 dokumentiert. Da der Messwert vor Bearbeitung der Bildgalerie bei der 2. Messung von Fall 12 Null war, kann dort keine prozentuale Abweichung berechnet werden (Division durch Null).

Fall	Messreihe	CC-Wert vor Bearbeitung der Bildgalerie	CC-Wert nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
1	1	0.40	-0.37	0.77	192.50
1	2	0.59	-0.37	0.96	162.71
2	1	0.19	-0.39	0.58	305.26
2	2	0.12	-0.37	0.49	408.33
3	1	0.02	-0.39	0.41	2050.00
3	2	0.04	-0.39	0.43	1075.00
4	1	-0.33	0.12	-0.45	136.36
4	2	-0.35	0.18	-0.53	151.43
5	1	0.23	-0.27	0.50	217.39
5	2	0.26	-0.15	0.41	157.69
6	1	-0.39	-0.38	-0,01	2.56
6	2	-0.09	-0.11	0.02	-22.22
7	1	0.49	-0.32	0.81	165.31

7	2	0.42	-0.30	0.72	171.43
8	1	0.07	-0.16	0.23	328.57
8	2	0.06	-0.26	0.32	533.33
9	1	-0.27	-0.17	-0.10	37.04
9	2	0.09	-0.23	0.32	355.56
10	1	0.23	-0.33	0.56	243.48
10	2	0.29	-0.39	0.68	234.48
11	1	0.02	-0.35	0.37	1850.00
11	2	-0.08	-0.18	0.10	-125.00
12	1	-0.11	-0.13	0.02	-18.18
12	2	0.00	-0.19	0.19	---
13	1	0.17	-0.29	0.46	270.59
13	2	0.15	-0.31	0.46	306.67
14	1	0.37	-0.08	0.45	121.62
14	2	0.54	-0.33	0.87	161.11
15	1	-0.35	-0.10	-0.25	71.43
15	2	-0.18	-0.06	-0.12	66.67
16	1	-0.15	-0.32	0.17	-113.33
16	2	0.07	-0.36	0.43	614.29
17	1	0.35	-0.10	0.45	128.57
17	2	-0.50	-0.39	-0.11	22.00
18	1	0.18	-0.38	0.56	311.11
18	2	0.34	-0.04	0.38	111.76
19	1	-0.08	-0.38	0.30	-375.00
19	2	-0.11	-0.39	0.28	-254.55
20	1	0.20	0.56	-0.36	-180.00
20	2	0.29	0.00	0.29	100.00
21	1	0.45	0.11	0.34	75.56
21	2	0.40	-0.05	0.45	112.50
22	1	0.31	0.02	0.29	93.55
22	2	-0.21	0.01	-0.22	104.76
23	1	0.11	-0.31	0.42	381.82
23	2	-0.14	-0.38	0.24	-171.43
24	1	0.28	0.00	0.28	100.00
24	2	0.33	-0.36	0.69	209.09
25	1	0.14	-0.31	0.45	321.43
25	2	0.30	-0.22	0.52	173.33
26	1	0.19	-0.19	0.38	200.00
26	2	-0.12	-0.39	0.27	-225.00

Tab. 12: Datenübersicht über die CC-Werte vor und nach Bearbeitung der Bildgalerie für beide Messreihen à 26 Messungen (n = 52)

Messreihe	CC-Wert vor Bearbeitung der Bildgalerie	CC-Wert nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
Beide Messreihen				
Mittelwert	0,10	-0,21	0,31	222,58
SD	0,26	0,19	0,32	421,23
CV	257,31	-91,20	104,42	189,25
Median	0,13	-0,28	0,375	157,69
Kleinster Wert	8,65	-0,39	-0,53	-375
Größter Wert	38,30	0,56	0,96	2050,
Messreihe 1				
Mittelwert	0,10	-0,19	0,29	266,06
SD	0,25	0,22	0,32	524,70
CV	239,24	-116,35	108,26	102,90
Median	0,175	-0,28	0,375	150,84
Kleinster Wert	-0,39	-0,39	-0,45	-375
Größter Wert	0,49	0,56	0,81	2050
Messreihe 2				
Mittelwert	0,10	-0,23	0,33	177,36
SD	0,27	0,16	0,34	280,52
CV	281,69	69,34	102,57	158,16
Median	0,08	-0,28	0,35	157,69
Kleinster Wert	-0,50	-0,39	-0,53	-254,55
Größter Wert	0,59	0,18	0,96	1075

Tab. 13: Statistische Grunddaten der CC-Werte vor und nach Bearbeitung der Bildgalerie. Angaben für beide Messreihen à 26 Messungen (n = 52) sowie für die einzelnen Messreihen 1 und 2 getrennt. SD = Standardabweichung, CV = Variationskoeffizient

Auch wenn die Werte allesamt niedrig wirken, zeigt sich doch eine beachtliche Standardabweichung, so dass die Werte letztendlich doch als recht breit gestreut anzusehen sind. Nach der Bearbeitung der Bildgalerie kamen sowohl höhere, als auch niedrigere Werte als vor der Bearbeitung der Bildgalerie vor, wobei sich die Werte insgesamt etwa harmonisch um Null herum verteilen.

Eine hohe Korrelation zwischen den Variablen der CC-Werte innerhalb der beiden Messreihen ließ sich insbesondere für den Vergleich der Werte vor bzw.

nach Bearbeitung der Bildgalerie nicht herstellen (0,0908 bzw. -0.1561) (Tabelle 14). Die Abnahmen der Werte hingegen („Differenz“) war mit den beiden Werten durchaus signifikant verbunden ($r > +/- 0,6$).

CC-Wert	vor Bearbeitung	nach Bearbeitung	Differenz	Differenz in %
Messreihe 1				
vor Bearbeitung	"---			
nach Bearbeitung	0.0908	"---		
Differenz	0.7250	-0.6201	"---	
Differenz in %	0.0120	-0.3325	0.2394	"---
Messreihe 2				
vor Bearbeitung	"---			
nach Bearbeitung	-0.1561	"---		
Differenz	0.8817	-0.6036	"---	
Differenz in %	0.0755	-0.0581	0..087	"---

Tab. 14: Korrelationskoeffizienten für die zum CC-Wert der Referenzzellen erhobenen Variablen, getrennt nach den Messreihen 1 und 2.

Im t-Test und im Mann-Whitney U-Test wiederum (Tabelle 15) traten keine Hinweise auf signifikante Unterschiede in der Datenverteilung zwischen den beiden Messreihen zutage ($p > 0,05$).

CC	t-Test	Mann-Whitney U-Test
vor Bearbeitung	0.9547	0.9025
nach Bearbeitung	0.4151	0.4559
Differenz	0.6630	0.7774
Differenz in %	0.4577	0.7992

Tab. 15: p-Werte des t-Tests und des Mann-Whitney U-Tests im Vergleich der Messreihen 1 und 2 für die zum CC-Wert der Referenzzellmessung erhobenen Variablen

4.3 Methodische Untersuchungen an den Analysezellen

Anders als bei den Referenzzellen wurde die statistische Auswertung an den Analysezellen nicht ausschließlich bezogen auf alle 26 Fälle vorgenommen, sondern aus inhaltlich-klinischen Gründen auch um Untersuchungen ergänzt, bei denen Daten für die beiden Gruppen Tribukait B und Tribukait C separat analysiert wurden. Dieses gilt sowohl für die Zahl der Analysezellen, als auch für den höchsten gemessenen Wert, den Maximalwert (C-Max). Im Folgenden werden sie als „abnorme Zellen“ bezeichnet, was der Zuordnung im Klassifikator als „Abnormals“ Rechnung trägt.

4.3.1 Zahl der abnormen Zellen

In Tabelle 16 sind die Daten für die 26 Messwertpaare bezogen auf die Zahl abnormer Zellen dargestellt, in Tabelle 17 die wichtigsten statistischen Kenndaten wie Mittelwert, Standardabweichung oder Median.

Fall	Messreihe	Zahl abnormer Zellen vor Bearbeitung der Bildgalerie	Zahl abnormer Zellen nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
1	1	17416	17155	261	1.50
1	2	14440	14389	51	0.35
2	1	5443	5420	23	0.42
2	2	5383	5358	25	0.46
3	1	22693	22542	151	0.67
3	2	22050	21993	57	0.26
4	1	13438	13295	143	1.06
4	2	15276	14880	396	2.59
5	1	14620	14460	160	1.09
5	2	14640	14468	172	1.17
6	1	4383	4127	256	5.84
6	2	4556	4459	97	2.13
7	1	5810	5684	126	2.17
7	2	5786	5732	54	0.93

8	1	13450	13202	248	1.84
8	2	12538	12268	270	2.15
9	1	10289	10205	84	0.82
9	2	12905	12752	153	1.19
10	1	1483	1373	110	7.42
10	2	1515	1365	150	9.90
11	1	2232	2070	162	7.26
11	2	1969	1820	149	7.57
12	1	1068	937	131	12.27
12	2	1222	1074	148	12.11
13	1	9019	8692	327	3.63
13	2	7738	7419	319	4.12
14	1	6398	6052	346	5.41
14	2	7669	7208	461	6.01
15	1	2094	1961	133	6.35
15	2	2394	2238	156	6.52
16	1	3839	3645	194	5.05
16	2	2404	2475	-71	-2.95
17	1	2773	2655	118	4.26
17	2	1280	1200	80	6.25
18	1	10742	10014	728	6.78
18	2	10367	9598	769	7.42
19	1	8190	8131	59	0.72
19	2	8455	8359	96	1.14
20	1	917	833	84	9.16
20	2	787	711	76	9.66
21	1	7850	7471	379	4.83
21	2	7435	7104	331	4.45
22	1	1375	1331	44	3.20
22	2	1577	1510	67	4.25
23	1	5483	5199	284	5.18
23	2	5434	5187	247	4.55
24	1	4585	4427	158	3.45
24	2	4240	4104	136	3.21
25	1	7160	6970	190	2.65
25	2	7223	7044	179	2.48
26	1	11443	9635	1808	15.80
26	2	10833	9804	1029	9.50

Tab. 16: Datenübersicht über die Zahl abnormer Zellen vor und nach Bearbeitung der Bildgalerie für beide Messreihen à 26 Messungen (n = 52)

Messreihe	Zahl abnormer Zellen vor Bearbeitung der Bildgalerie	Zahl abnormer Zellen nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
Beide Messreihen				
Mittelwert	7390,6	7153,9	236,6	4,35
SD	5463,7	5399,6	294,8	3,67
CV	73,9%	75,5%	124,6%	84,4%
Median	6104	5892	152	3,87
Kleinster Wert	787	711	-71	-2,95
Größter Wert	22693	22542	1808	15,80
Messreihe 1				
Mittelwert	7469,0	7211,0	258,0	4,57
SD	5546,4	5475,8	346,7	3,74
CV	74,3%	75,9%	134,4%	81,8%
Median	6104	5868	159	3,94
Kleinster Wert	917	833	23	0,42
Größter Wert	22693	22542	1808	15,80
Messreihe 2				
Mittelwert	7312,2	7096,9	215,3	4,13
SD	5488,5	5430,1	237,0	3,67
CV	75,1%	76,5%	110,1%	88,8%
Median	6504,5	6388	149,5	3,67
Kleinster Wert	787	711	-71	-2,95
Größter Wert	22050	21993	1029	12,11

Tab. 17: Statistische Grunddaten der Zahl abnormer Zellen vor und nach Bearbeitung der Bildgalerie. Angaben für beide Messreihen à 26 Messungen (n = 52) sowie für die einzelnen Messreihen 1 und 2 getrennt. SD = Standardabweichung, CV = Variationskoeffizient

Die Daten zeigen zum einen, wie unterschiedlich groß die Zahl der automatisiert als abnorm erkannten Messobjekte, die zwischen 787 und 22693 schwankt, ist, abhängig vom rein technischen Ergebnis der Zellvereinzelung, welches wiederum auf biologischen Gegebenheiten des Grundgewebes basiert. Außerdem lässt sich ablesen, dass sich die Tätigkeit des Untersuchers auf die durchschnittliche Reduktion der Zahl von Messobjekten um 4,35% +/- 3,74%

(SD) der Objekte beschränkt. Allerdings können von Fall zu Fall gehörige Schwankungen auftreten. Bei nur einer Messung kamen „netto“ Objekte hinzu (Abweichung: -2,95%).

Nicht dokumentiert wurde, wie viele Objekte insgesamt ausgetauscht wurden, also ob eine Reduktion von z.B. 30 Objekten rein durch das Elimination von 30 Objekten oder durch eine Elimination von z.B. 50 Objekten und die Hinzu- nahme 20 anderer, fälschlich als Artefakte kategorisierter Objekte zustande kam.

Die Korrelationskoeffizienten für die Variablen in den einzelnen Messreihen sind in Tabelle 18 wiedergegeben. Insbesondere für die Objektzahl vor und nach Bearbeitung der Bildgalerie ergaben sich sehr hohe Werte (0,9981 bzw. 0,9991).

Abnorme Zellen	vor Bearbeitung	nach Bearbeitung	Differenz	Differenz in %
	Messreihe 1			
vor Bearbeitung	"---			
nach Bearbeitung	0.9981	"---		
Differenz	0.2339	0.1736	"---	
Differenz in %	-0.4184	-0.4634	0.6256	"---
	Messreihe 2			
vor Bearbeitung	"---			
nach Bearbeitung	0.9991	"---		
Differenz	0.2605	0.2186	"---	
Differenz in %	-0.3987	-0.4261	0.5158	"---

Tab. 18: Korrelationskoeffizienten für die zu der Zahl der abnormen Zellen erhobenen Variablen, getrennt nach den Messreihen 1 und 2.

Im Vergleich beider Messreihen zeigte sich weder für den t-Test, noch für den Mann-Whitney U-Test eine statistische Signifikanz ($p > 0,05$), was auf eine gute Vergleichbarkeit beider Messreihen schließen lässt (Tabelle 19).

4.3.2 Höhe des maximalen Messwertes (C-Max)

In den Tabellen 20 und 21 sind die Messwerte und die statistischen Kenndaten für den maximalen Messwert dargestellt.

Dabei zeigte sich, dass die Mittelwerte und Mediane zwar gut vergleichbar waren, dass aber andererseits doch deutliche Abweichungen zwischen den beiden Messreihen zutage traten. Für die meisten Messungen ergab sich im Rahmen der Durchsicht der Bildgalerien, dass besonders hohe Einzelwerte als Überlappungen von Zellen, z.B. als Dubletten, erkannt und eliminiert wurden. Dafür wurden allerdings vom Untersucher andere Messobjekte, die als Artefakte eingeordnet worden waren, als Messzellen akzeptiert. Insgesamt waren am Ende die C-Max-Werte in 28 Messungen nach Durchsicht der Bildgalerien für Messobjekte und Artefakte höher und in 24 Messungen niedriger als zuvor, was in etwa ausgeglichen ist.

Fall	Messreihe	C-Max vor Bearbeitung der Bildgalerie	C-Max nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
1	1	23.17	9.71	13.46	58.09
1	2	6.65	9.44	-2.79	-41.95
2	1	4.56	4.47	0.09	1.97
2	2	4.69	4.78	-0.09	-1.92
3	1	7.97	7.17	0.80	10.04
3	2	7.98	9.01	-1.03	-12.91
4	1	9.26	10.64	-1.38	-14.90
4	2	8.74	16.40	-7.66	-87.64
5	1	8.76	8.38	0.38	4.34
5	2	9.07	7.25	1.82	20.07
6	1	9.47	10.33	-0.86	-9.08
6	2	9.51	10.56	-1.05	-11.04
7	1	7.50	5.49	2.01	26.80
7	2	5.16	5.34	-0.18	-3.49
8	1	5.88	5.89	-0,01	-0.17
8	2	8.41	5.10	3.31	39.36
9	1	8.52	6.86	1.66	19.48

9	2	8.51	5.03	3.48	40.89
10	1	9.18	5.67	3.51	38.24
10	2	8.89	5.22	3.67	41.28
11	1	11.30	6.33	4.97	43.98
11	2	11.71	5.23	6.48	55.34
12	1	4.96	8.23	-3.27	-65.93
12	2	5.66	7.84	-2.18	-38.52
13	1	15.47	12.26	3.21	20.75
13	2	8.98	9.77	-0.79	-8.80
14	1	3.85	7.41	-3.56	-92.47
14	2	4.44	7.62	-3.18	-71.62
15	1	5.50	8.58	-3.08	-56.00
15	2	5.50	7.78	-2.28	-41.45
16	1	5.05	6.04	-0.99	-19.60
16	2	7.26	5.61	1.65	22.73
17	1	5.45	6.87	-1.42	-26.06
17	2	5.09	6.61	-1.52	-29.86
18	1	11.67	11.91	-0.24	-2.06
18	2	21.55	9.94	11.61	53.87
19	1	6.85	30.64	-23.79	-347.30
19	2	9.71	5.77	3.94	40.58
20	1	8.50	17.11	-8.61	-101.29
20	2	8.10	19.64	-11.54	-142.47
21	1	11.26	18.73	-7.47	-66.34
21	2	10.34	19.70	-9.36	-90.52
22	1	9.28	14.69	-5.41	-58.30
22	2	10.25	11.36	-1.11	-10.83
23	1	6.75	6.64	0.11	1.63
23	2	8.34	5.32	3.02	36.21
24	1	11.36	6.04	5.32	46.83
24	2	7.88	6.47	1.41	17.89
25	1	6.13	4.74	1.39	22.68
25	2	6.26	4.79	1.47	23.48
26	1	38.44	10.75	27.69	72.03
26	2	8.39	9.21	-0.82	-9.77

Tab. 20: Datenübersicht über den höchsten Messwert (C-Max) vor und nach Bearbeitung der Bildgalerie für beide Messreihen à 26 Messungen (n = 52)

Messreihe	C-Max vor Bearbeitung der Bildgalerie	C-Max nach Bearbeitung der Bildgalerie	Differenz	Differenz in %
Beide Messreihen				
Mittelwert	9,09	9,08	0,01	-13,53
SD	5,51	4,97	6,80	66,61
CV	60,57	54,76	44769,68	-492,18
Median	8,40	7,515	-0,135	-1,99
Kleinsten Wert	3,85	4,47	-23,79	-347,3
Größten Wert	38,44	30,64	27,69	72,0
Messreihe 1				
Mittelwert	9,85	9,68	0,17	-18,95
SD	7,07	5,63	8,48	80,76
CV	71,75	58,26	48,86	-4,26
Median	8,51	7,82	0,04	0,73
Kleinsten Wert	3,85	4,47	-23,79	-347,30
Größten Wert	38,44	30,64	27,69	72,03
Messreihe 2				
Mittelwert	8,35	8,49	-0,14	-8,11
SD	3,29	4,24	4,74	49,67
CV	39,46	49,92	-3313,70	-611,76
Median	8,375	7,435	-0,485	-6,14
Kleinsten Wert	4,44	4,78	-11,54	-142,47
Größten Wert	21,55	19,70	11,61	55,34

Tab. 21: Statistische Grunddaten über den höchsten Messwert (C-Max) vor und nach Bearbeitung der Bildgalerie. Angaben für beide Messreihen à 26 Messungen (n = 52) sowie für die einzelnen Messreihen 1 und 2 getrennt. SD = Standardabweichung, CV = Variationskoeffizient

Aufgrund der doch recht deutlichen Abweichungen der Messwerte zueinander bestand keine eindeutige Korrelation zwischen den C-Max-Werten vor und nach Bearbeitung der Bildgalerie (0,1240 bzw. 0,2253) (Tabelle 22).

C-Max	vor Bearbeitung	nach Bearbeitung	Differenz	Differenz in %
	Messreihe 1			
vor Bearbeitung	---			
nach Bearbeitung	0.1240	---		
Differenz	0.7513	-0.5617	---	
Differenz in %	0.3717	-0.7998	0.8418	---
	Messreihe 2			
vor Bearbeitung	---			
nach Bearbeitung	0.2253	---		
Differenz	0.4858	-0.7422	---	
Differenz in %	0.3261	-0.8037	0.9453	---

Tab. 22: Korrelationskoeffizienten für die zum maximalen Messwert (C-Max) erhobenen Variablen, getrennt nach den Messreihen 1 und 2.

Abschließend sind noch die Ergebnisse des t-Tests bzw. des Mann-Whitney U-Test wiedergegeben (Tabelle 23). Bei einem p-Wert von jeweils $> 0,05$ lässt sich annehmen, dass die Messwertpaare beider Messreihen eine gute Vergleichbarkeit untereinander aufweisen.

C-Max	t-Test	Mann-Whitney U-Test
vor Bearbeitung	0.3750	0.7559
nach Bearbeitung	0.4157	0.3000
Differenz	0.9041	0.9100
Differenz in %	0.5276	0.8211

Tab. 23: p-Werte des t-Tests und des Mann-Whitney U-Tests im Vergleich der Messreihen 1 und 2 für die zum maximalen Messwert (C-Max) erhobenen Variablen

Für alle 52 Messungen gibt Abb. 25 die Beziehung zwischen dem maximalen Messwert (C-Max) vor und nach Bearbeitung der Bildgalerie wieder. Dabei zeigt sich bis zu einem C-Max-Wert von ca. 15c eine vorwiegend als linear zu bezeichnende Beziehung der Messwerte.

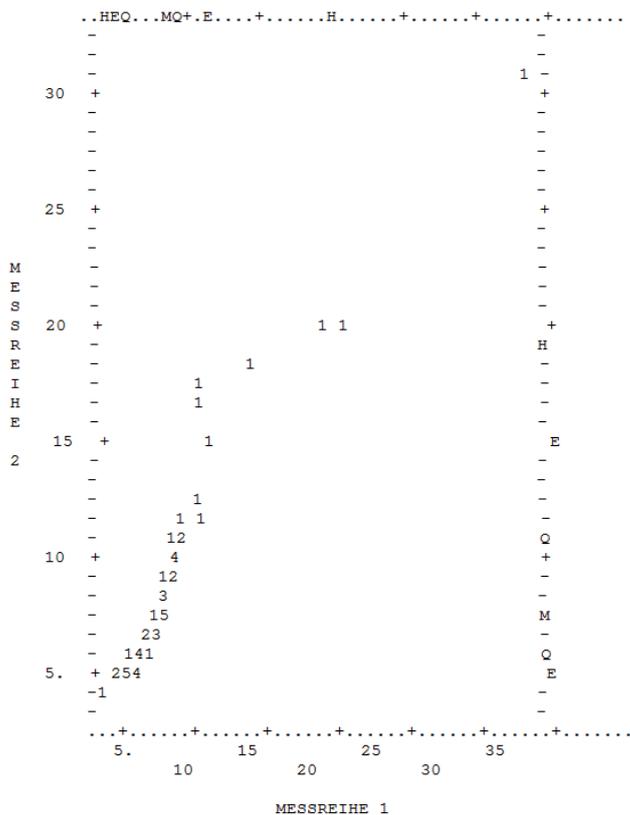


Abb. 25: Graphische Darstellung der 52 Messwerte der Messreihen 1 und 2 für den maximalen Messwert (C-Max) vor und nach Bearbeitung der Bildgalerie. Die Ziffern 1 bis 5 in der Grafik geben wieder, wie viele Messpunkte auf einen Datenpunkt fallen. Die Ziffernsumme beträgt entsprechend 52.

4.3.3 Lage des ersten Peak

Die Lage des ersten Peaks ist zwar formal ein sich kontinuierlich erstreckender Parameter, wird aber mit bestimmten festen Grenzen interpretiert (siehe Material und Methoden). Deswegen macht es Sinn, seine Datenverteilung genauer zu betrachten und auch natürliche Schwankungsbreiten zu ermitteln, mit denen man von Messung zu Messung rechnen muss. Die entsprechenden Berechnungen wurden aus klinischen Gründen für die als Tribukait B bzw. als Tribukait C beurteilten Fälle getrennt vorgenommen.

4.3.3.1 Tribukait B-Fälle

Vergleich zwischen manueller Messung und automatisierter Messung

Der im Folgenden wiedergegebene Vergleich berücksichtigt die manuelle Messung und als repräsentatives Beispiel für die Automatisierung das Ergebnis der ersten automatisierten Messung. Tabelle 24 stellt die Ergebnisse der Messungen gegenüber. Der durchschnittliche Wert für die automatisierten Messungen liegt bei 2,0325c und für die manuellen Messungen bei 2,015c, die mittlere Differenz bei 0,0175c. Die größte Differenz zwischen zwei Messungen lag bei 0,09c. Alle Messungen lagen im Toleranzbereich von $2c \pm 10\%$, so dass keine aneuploiden DNA-Stammlinien zu beobachten waren. Dass von 12 Messungen neun automatisiert höher lagen und nur drei niedriger als bei der manuellen Messung, kann bei den minimalen Abweichungen vernachlässigt werden und ist hier nicht als relevanter oder gar klinisch entscheidender Trend für eine Neigung der automatisierten Messungen zu höheren Messwerten zu interpretieren.

Tribukait B			
Fall	Manuelle Messung	Automatisierte Messung	Differenz
1	1,98c	2,01c	0,03c
2	2,04c	2,05c	0,01c
3	2,01c	2,06c	0,05c
4	2,02c	2,00c	-0,02c
5	2,05c	2,07c	0,02c
6	1,99c	2,00c	0,01c
7	2,04c	2,02c	-0,02c
8	2,00c	2,03c	0,03c
9	2,04c	2,03c	-0,01c
10	2,01c	2,02c	0,01c
11	2,00c	2,09c	0,09c
12	2,00c	2,01c	0,01c
	Ø 2,015c	Ø 2,0325c	Ø 0,0175c

Tab. 24: Datenübersicht zur Lage des ersten Peaks im Vergleich der manuellen und der automatisierten Messung für Tribukait B-Fälle

Die Tests auf Normalverteilung zeigten, dass die Daten normalverteilt sind und somit die parametrische Statistik anwendbar war. In Tabelle 25 sind die entsprechenden Daten dargestellt.

Spearman's Korrelationskoeffizient		t-Test der Stichproben		Konfidenzintervall des Bias		Konfidenzintervall untere Grenze der Übereinstimmung (95% Konfidenz)		Konfidenzintervall obere Grenze der Übereinstimmung (95% Konfidenz)	
Spearman's Rho	p-wert	t	p-wert	untere	obere	untere	obere	untere	obere
0,166	0,606	1,969	0,075	-0,0021	0,0371	-0,0763	-0,0094	0,0444	0,1113

Tab. 25: Teststatistik zum Vergleich der Lage des Peak 1 im Vergleich automatischer und manueller Messungen für Tribukait B-Fälle.

Aus den Daten und der Wertetabelle oben kann man ablesen, dass die Messwerte des Peak 1 der Diagnosegruppe Tribukait B weitestgehend übereinstimmend sind. Auch der Bland-Altman-Plot (Abb. 26) zeigt keine auffälligen Schwankungen der Messwerte bei einer maximalen Abweichung von 0,09c und lässt erkennen, dass die gemessenen Schwankungen unabhängig von der Lage der Messwerte sind.

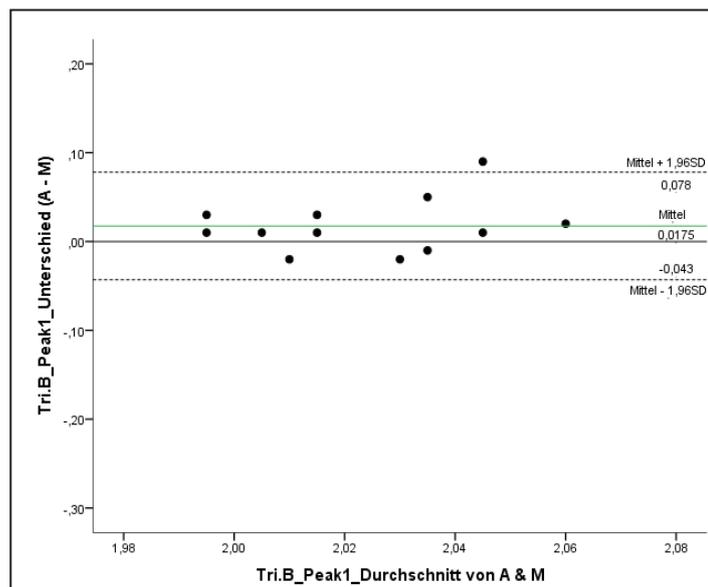


Abb. 26: Bland-Altman-Plot für Peak 1 im Vergleich der manuellen Messung (M) und der ersten automatisierten Messung (A) der Tribukait B-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

Vergleich zwischen den beiden automatisierten Messungen

Die Daten für den Vergleich der beiden automatisierten Messungen sind in Tabelle 26 wiedergegeben. Auch hier stellen sich die Messwerte weitgehend gleichlautend für beide Messungen dar. Der durchschnittliche Wert für die Messungen liegt bei 2,0325c bzw. bei 2,0367c und unterscheidet sich somit um weniger als 0,01c. Die größte Differenz zwischen zwei Messungen betrug 0,09c. Alle Messwerte verblieben im Toleranzbereich von $2c \pm 10\%$, so dass keine aneuploiden DNA-Stammlinien auftraten. In jeweils sechs Fällen lagen die erste oder die zweite Messung um einen geringen Betrag höher bzw. niedriger, so dass sich hier kein auswertbarer Trend darstellte.

Tribukait B			
Fall	Automatisierte Messung 1	Automatisierte Messung 2	Differenz
1	2,01c	2,02c	-0,01c
2	2,05c	2,06c	-0,01c
3	2,06c	2,01c	0,05c
4	2,00c	2,00c	0,00c
5	2,07c	2,05c	0,02c
6	2,00c	1,97c	0,03c
7	2,02c	2,00c	0,02c
8	2,03c	1,95c	0,08c
9	2,03c	2,08c	-0,05c
10	2,02c	2,07c	-0,05c
11	2,09c	2,13c	-0,04c
12	2,01c	2,10c	-0,09c
	Ø 2,0325c	Ø 2,0367c	Ø 0,0042c

Tab. 26: Datenübersicht zur Lage des ersten Peaks im Vergleich der beiden automatisierten Messungen für die Tribukait B-Fälle

Die Tests auf Normalverteilung zeigten, dass auch diese Daten normalverteilt waren und somit die parametrische Statistik angewendet werden konnte. In Tabelle 27 sind die entsprechenden Daten dargestellt.

Spearman's Korrelationskoeffizient		t-Test der Stichproben (Bias null)		Konfidenzintervall des Bias		Konfidenzintervall untere Grenze der Übereinstimmung (95% Konfidenz)		Konfidenzintervall obere Grenze der Übereinstimmung (95% Konfidenz)	
<i>Spearman's Rho</i>	<i>p</i> -wert	t	<i>p</i> -wert	untere	obere	untere	obere	untere	obere
-0,568	0,054	-0,267	0,794	-0,0618	0,0132	-0,0700	-0,0188	0,0224	0,0737

Tab. 27: Teststatistik zum Vergleich der Lage des Peak 1 im Vergleich der beiden automatisierten Messungen für Tribukait B-Fälle.

Aus den Daten und der Wertetabelle oben kann man ablesen, dass die Messwerte des Peak 1 der Diagnosegruppe Tribukait B weitestgehend übereinstimmen. Auch der Bland-Altman-Plot (Abb. 27) zeigt keine auffälligen Schwankungen der Messwerte bei einer maximalen Abweichung von 0,09c.

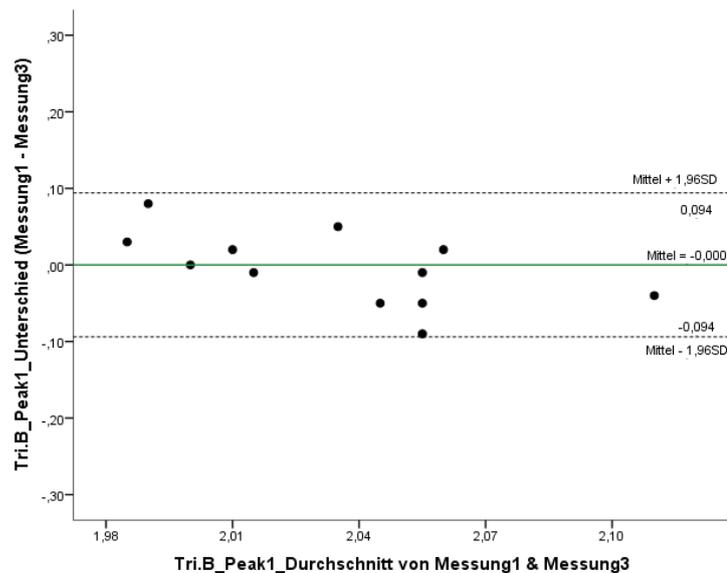


Abb. 27: Bland-Altman-Plot für Peak 1 im Vergleich der beiden automatisierten Messungen der Tribukait B-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

4.3.3.2 Tribukait C-Fälle

Vergleich zwischen manueller Messung und automatisierter Messung

Tabelle 28 zeigt die Messwerte für den Vergleich zwischen der manuellen Messung und der ersten automatisierten Messung. Der durchschnittliche Wert

für die automatisierten Messungen beträgt 2,08c, für die manuellen Messungen 2,045c, die mittlere Differenz 0,034c. Die größte Differenz zwischen zwei Messungen lag in zwei Fällen bei 0,15c. Vier manuelle Messwerte lagen am Rand des Toleranzbereiches von $2c \pm 10\%$, wurden aber als noch nicht Beleg für eine aneuploide DNA-Stammlinie interpretiert.

Dass von 15 Messungen zwölf automatisiert höher lagen und nur drei niedriger als bei der manuellen Messung, ist zwar ein zu beachtender Trend, spielt aber bei den im Grunde sehr geringen Abweichungen keine für klinische oder methodische Aussagen wesentliche Rolle.

Tribukait C			
Fall	Manuelle Messung	Automatisierte Messung	Differenz
1	2,00c	2,11c	-0,11c
2	2,04c	2,17c	-0,13c
3	2,21c	2,06c	0,15c
4	2,20c	2,09c	0,11c
5	2,20c	2,05c	0,15c
6	2,06c	2,00c	0,06c
7	1,99c	1,98c	0,01c
8	1,95c	1,91c	0,04c
9	2,10c	2,09c	0,01c
10	2,05c	2,05c	0,00c
11	2,13c	2,07c	0,06c
12	2,06c	1,95c	0,11c
13	2,00c	2,00c	0,00c
14	2,20c	2,06c	0,14c
15	2,01c	2,08c	-0,07c
	Ø 2,08c	Ø 2,045c	Ø 0,034c

Tab. 28: Datenübersicht zur Lage des ersten Peaks im Vergleich der manuellen und der automatisierten Messung für Tribukait C-Fälle

Die Tests auf Normalverteilung zeigten, dass die Daten normalverteilt sind und somit die parametrische Statistik anwendbar war. In Tabelle 29 sind die entsprechenden Daten dargestellt.

Spearman's Korrelationskoeffizient		t-Test der Stichproben		Konfidenzintervall des Bias		Konfidenzintervall untere Grenze der Übereinstimmung (95% Konfidenz)		Konfidenzintervall obere Grenze der Übereinstimmung (95% Konfidenz)	
Spearman's Rho	p-wert	t	p-wert	untere	obere	untere	obere	untere	obere
-0,445	0,096	-1,523	0,150	-0,0851	0,0144	-0,2965	-0,1263	0,0557	0,2258

Tab. 29: Teststatistik zum Vergleich der Lage des Peak 1 im Vergleich automatischer und manueller Messungen für Tribukait C-Fälle.

Aus den Daten und der Wertetabelle oben kann man ablesen, dass die Messwerte des Peak 1 der Diagnosegruppe Tribukait C eine gute Übereinstimmung zeigen. Der Bland-Altman-Plot (Abb. 28) zeigt alle Wertepaare im Toleranzbereich, außerdem lässt sich ablesen, dass die gemessenen Schwankungen offenbar als unabhängig von der Lage der Messwerte betrachtet werden können.

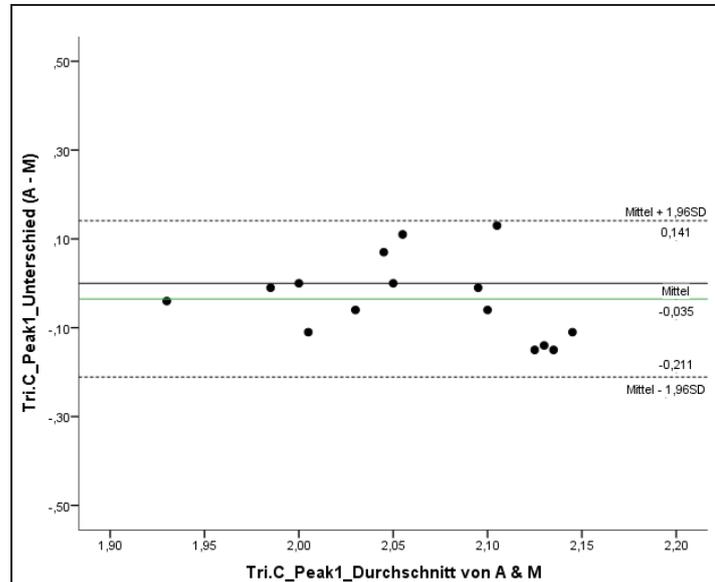


Abb. 28: Bland-Altman-Plot für Peak 1 im Vergleich der manuellen Messung (M) und der ersten automatisierten Messung (A) der Tribukait C-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

Vergleich zwischen den beiden automatisierten Messungen

Die Daten für den Vergleich der beiden automatisierten Messungen zeigt Tabelle 30. Bei Fall 13 ließ sich die zweite Messung nicht verlässlich auswerten, so dass der Fall hier nicht berücksichtigt werden konnte. Insgesamt stellen sich die Messwerte weitgehend gleichlautend für beide Messungen dar; die durchschnittlichen Werte für die Messungen liegen bei 2,045c bzw. bei 2,016c und unterscheiden sich somit nur um fast genau 0,03c. Die größte Differenz zwischen zwei Messungen betrug 0,08c. Alle Messwerte lagen bei $2c \pm 10\%$ und somit im Toleranzbereich von; aneuploiden DNA-Stammlinien kamen nicht vor. Bei neun Fällen lag der erste Messwert höher, bei fünf den zweiten, was bei den geringen Unterschieden nicht für einen auswertbaren Trend spricht.

Tribukait C			
Fall	Automatisierte Messung 1	Automatisierte Messung 2	Differenz
1	2,11c	2,03c	0,08c
2	2,17c	2,10c	0,07c
3	2,06c	2,00c	0,06c
4	2,09c	2,15c	-0,06c
5	2,05c	2,00c	0,05c
6	2,00c	1,95c	0,05c
7	1,98c	2,05c	-0,07c
8	1,91c	1,95c	-0,04c
9	2,09c	2,03c	0,06c
10	2,05c	2,06c	-0,01c
11	2,07c	2,09c	-0,02c
12	1,95c	1,83c	0,12c
13	2,00c	-	-
14	2,06c	1,94c	0,12c
15	2,08c	2,04c	0,04c
	Ø 2,045c	Ø 2,016c	Ø 0,032c

Tab. 30: Datenübersicht zur Lage des ersten Peaks im Vergleich der beiden automatisierten Messung für die Tribukait C-Fälle

Die Tests auf Normalverteilung zeigten, dass auch an diesen Daten eine parametrische Statistik angewendet werden konnte. In Tabelle 31 sind die entsprechenden Daten wiedergegeben.

Spearman's Korrelationskoeffizient		t-Test der Stichproben		Konfidenzintervall des Bias		Konfidenzintervall untere Grenze der Übereinstimmung (95% Konfidenz)		Konfidenzintervall obere Grenze der Übereinstimmung (95% Konfidenz)	
Spearman's Rho	p-wert	t	p-wert	untere	obere	untere	obere	untere	obere
-0,130	0,658	1,939	0,075	-0,0037	0,0680	-0,1507	-0,0282	0,0925	0,2150

Tab. 31: Teststatistik zum Vergleich der Lage des Peak 1 im Vergleich der beiden automatisierten Messungen für Tribukait C-Fälle.

Aus den Daten und der Wertetabelle oben kann man ablesen, dass die Messwerte des Peak 1 der Diagnosegruppe Tribukait C weitestgehend übereinstimmend sind. Der Bland-Altman-Plot (Abb. 29) weist keine auffälligen Schwankungen der Messwerte auf; die Breite des Konfidenzintervalls des Vertrauensbereiches wird weder ausgeschöpft, noch überschritten. Eine Abhängigkeit der gemessenen Unterschiede von der Höhe der Messwerte ist nicht zu erkennen.

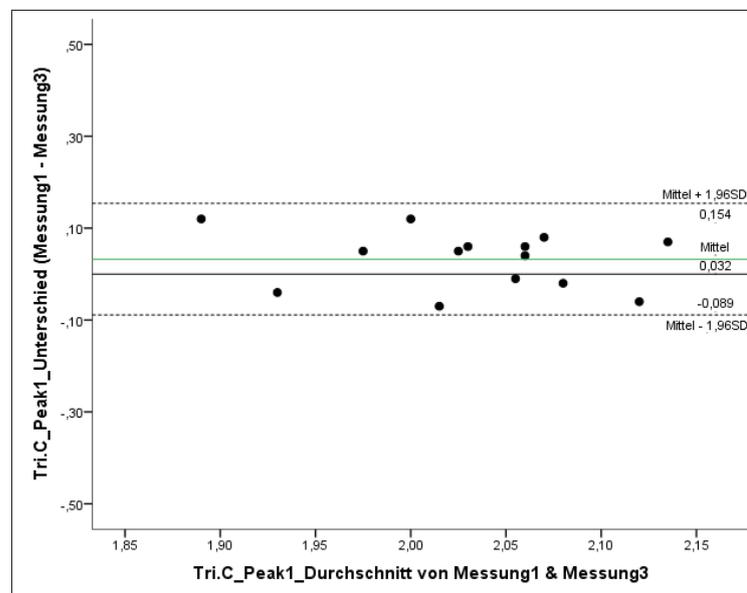


Abb. 29: Bland-Altman-Plot für Peak 1 im Vergleich der beiden automatisierten Messungen der Tribukait C-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

4.3.4 Lage des zweiten Peak

4.3.4.1 Tribukait B-Fälle

Vergleich zwischen manueller Messung und automatisierter Messung

Der im Folgenden wiedergegebene Vergleich bezieht sich auf die manuelle Messung und als repräsentatives Beispiel das Ergebnis der ersten automatisierten Messung. Tabelle 32 stellt die Messergebnisse dar. Der durchschnittliche Wert für die automatisierten Messungen liegt bei 4,07c und für die manuellen Messungen bei 4,25c, die mittlere Differenz bei -0,18c, die größte Differenz bei -0,67c. Nicht alle automatisierten Messungen hielten den Toleranzbereich von $4c \pm 10\%$ ein; eine der 12 Messungen lag mit 4,70c klar über der Schwelle von 4,4c, eine weitere ergab gerade noch den Schwellenwert. Dass von 12 Messungen elf automatisiert höher lagen als bei der manuellen Messung, bestätigt den gleichartigen Trend beim ersten Peak (dort 9:3). Auch wenn die durchschnittliche Abweichung mit -0,18c nicht sehr hoch wirkt, beträgt sie doch bereits 45% und damit fast die Hälfte der zu einer Seite von 4c erlaubten Abweichung bei der klinischen Interpretation (0,18c von 0,40c).

Tribukait B			
Fall	Manuelle Messung	Automatisierte Messung	Differenz
1	3,97c	4,26c	-0,29c
2	4,29c	4,31c	-0,02c
3	4,04c	4,21c	-0,17c
4	4,03c	4,70c	-0,67c
5	4,27c	4,38c	-0,11c
6	3,98c	4,14c	-0,16c
7	3,98c	4,04c	-0,06c
8	4,00c	4,14c	-0,14c
9	4,15c	4,31c	-0,16c
10	4,01c	4,41c	-0,40c
11	4,00c	3,90c	0,10c
12	4,15c	4,20c	-0,05c
	Ø 4,07c	Ø 4,25c	Ø -0,18c

Tab. 32: Datenübersicht zur Lage des zweiten Peaks im Vergleich der manuellen und der automatisierten Messung für Tribukait B-Fälle

Die Tests auf Normalverteilung zeigten, dass die Daten normalverteilt sind und somit die parametrische Statistik anwendbar war. In Tabelle 33 sind die entsprechenden Daten dargestellt.

Spearman Korrelationskoeffizient		Konfidenzintervall (95% Konfidenz) der Schätzung 75% ($\pm 0,2$)		Konfidenzintervall (95% Konfidenz) der Schätzung 83% ($\pm 0,3$)	
<i>Spearman-Rho</i>	<i>p</i> -wert	untere	obere	untere	Obere
0,315	0,319	42,8%	94,5%	51,6%	97,9%

Tab. 33: Teststatistik zum Vergleich der Lage des Peak 2 im Vergleich automatischer und manueller Messungen für Tribukait B-Fälle.

Aus den Daten und der Wertetabelle oben kann man ablesen, dass die Messwerte des Peak 2 der Diagnosegruppe Tribukait B vergleichbar sind, aber im Trend bei der automatisierten Messung höher liegen. Auch der Bland-Altman-Plot (Abb. 30) illustriert diesen Sachverhalt noch einmal gut, erkennbar daran, dass die meisten Werte oberhalb der Null-Linie der y-Achse liegen. Zwei Fälle liegen außerhalb des Vertrauensbereiches, einer davon deutlich (Fall 4).

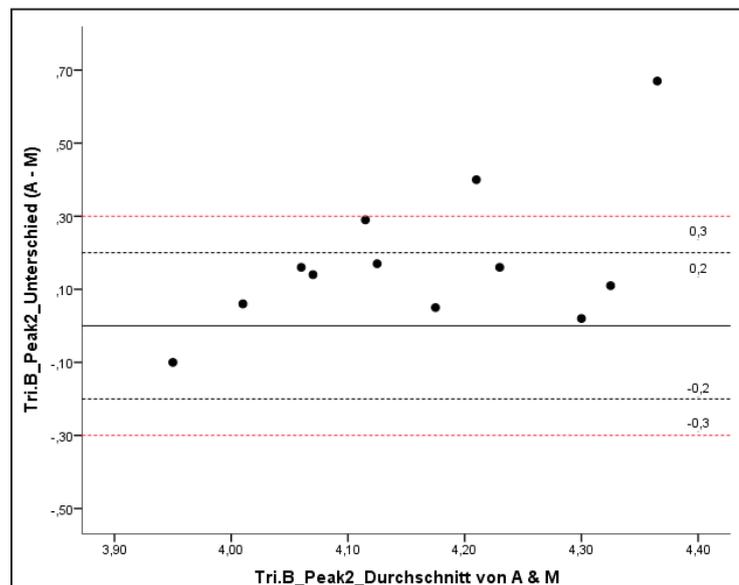


Abb. 30: Bland-Altman-Plot für Peak 2 im Vergleich der manuellen Messung (M) und der ersten automatisierten Messung (A) der Tribukait B-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

Vergleich zwischen den beiden automatisierten Messungen

Die Daten für den Vergleich der beiden automatisierten Messungen sind in Tabelle 34 wiedergegeben. Zu bemerken ist, dass sich in drei Fällen bei der zweiten Messung kein eigentlicher Peak abgrenzen ließ. Für die neun verbliebenen Fälle stellen sich die Messwerte annähernd gleich für beide Messungen dar. Die Mittelwerte für die Messungen betragen 4,26c bzw. 4,25c und unterscheiden sich somit lediglich um 0,01c. Die größte Differenz zwischen zwei Messungen betrug allerdings doch bereits 0,24c. Acht Fälle zeigten Messwerte innerhalb des Toleranzbereiches von $4c \pm 10\%$, im neunten Fall lagen beide Peaks mit fast gleichen Werten von 4,70c und 4,60c jeweils außerhalb. In fünf Fällen zeigte die erste Messung einen höheren Wert für den zweiten Peak, in vier Fällen die zweite, was ausgeglichen ist.

Tribukait B			
Fall	Automatisierte Messung 1	Automatisierte Messung 2	Differenz
1	4,26c	4,30c	-0,04c
2	4,31c	4,39c	-0,08c
3	4,21c	4,31c	-0,10c
4	4,70c	4,69c	0,01c
5	4,38c	4,14c	0,24c
6	4,14c	4,06c	0,08c
7	4,04c	4,20c	-0,16c
8	4,14c	4,11c	0,03c
9	4,31c	-	-
10	4,41c	-	-
11	3,90c	-	-
12	4,20c	4,09c	0,11c
	Ø 4,26c	Ø 4,25c	Ø 0,01c

Tab. 34: Datenübersicht zur Lage des zweiten Peaks im Vergleich der beiden automatisierten Messungen für die Tribukait B-Fälle. Im Mittelwert der automatisierten Messung 1 sind die Messwerte der Fälle 9-11 nicht berücksichtigt.

Die Tests auf Normalverteilung zeigten, dass auch diese Daten normalverteilt waren und somit die parametrische Statistik angewendet werden konnte. In Tabelle 35 sind die entsprechenden Daten dargestellt.

Spearman's Korrelationskoeffizient		t-Test der Stichproben		Konfidenzintervall des Bias		Konfidenzintervall untere Grenze der Übereinstimmung (95% Konfidenz)		Konfidenzintervall obere Grenze der Übereinstimmung (95% Konfidenz)	
Spearman's Rho	p-Wert	t	p-Wert	untere	obere	untere	obere	untere	obere
-0,142	0,715	0,245	0,813	-0,0841	0,1041	-0,3907	-0,0690	0,0890	0,4107

Tab. 35: Teststatistik zum Vergleich der Lage des Peak 2 im Vergleich der beiden automatisierten Messungen für Tribukait B-Fälle.

Aus den Daten und der Wertetabelle ergibt sich, dass die Messwerte des Peak 2 der Diagnosegruppe Tribukait B bei den beiden automatisierten Messungen zwar in den Mittelwerten fast genau übereinstimmen, aber dennoch eine gewisse Schwankung der Werte vorhanden ist. Der Bland-Altman-Plot (Abb. 31) illustriert diesen Sachverhalt. Dabei bleiben alle Wertepaare im Vertrauensbereich, und eine Abhängigkeit der gemessenen Unterschiede von der Höhe der Messwerte ist nicht erkennbar.

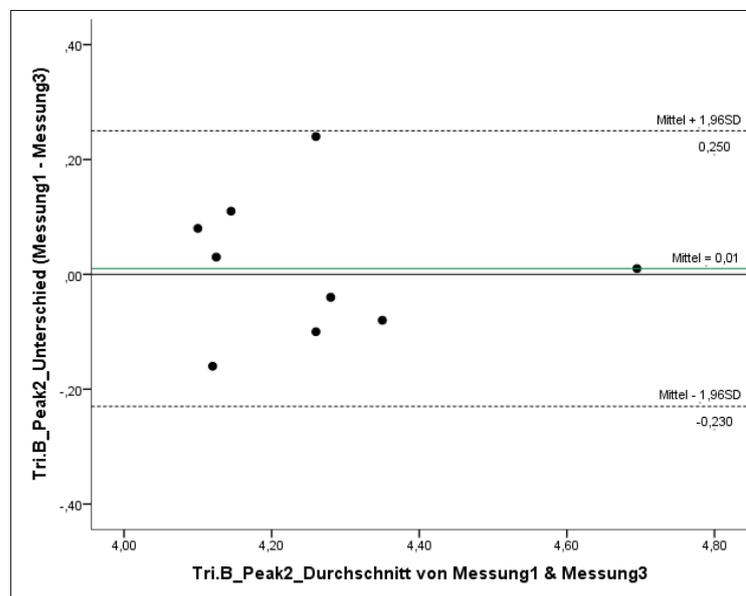


Abb. 31: Bland-Altman-Plot für Peak 2 im Vergleich der beiden automatisierten Messungen der Tribukait B-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

4.3.4.2 Tribukait C-Fälle

Vergleich zwischen manueller Messung und automatisierter Messung

Tabelle 36 stellt die Messwerte für den Vergleich zwischen der manuellen Messung und der ersten automatisierten Messung dar. Der durchschnittliche Wert für die automatisierten Messungen lag bei 4,17c, für die manuellen Messungen bei 4,38c; die mittlere Differenz betrug -0,21c. Die größte Differenz zwischen zwei Messungen lag bei -0,68c. Fünf manuelle und sieben automatisierte Messwerte befanden sich außerhalb des Toleranzbereiches von 4c +/- 10% und wurden als Beleg für eine aneuploide DNA-Stammlinie interpretiert. Dass von 15 Messungen elf automatisiert höhere Werte zeigten und nur vier niedriger lagen als bei der manuellen Messung, ist ein zu beachtender Trend, spielt aber bei den noch recht geringen Abweichungen keine wesentliche Rolle.

Tribukait C			
Fall	Manuelle Messung	Automatisierte Messung	Differenz
1	4,10c	4,38c	-0,28c
2	3,43c	3,90c	-0,47c
3	4,47c	4,44c	0,03c
4	4,30c	3,92c	0,38c
5	4,30c	4,20c	0,10c
6	4,03c	4,29c	-0,26c
7	3,09c	3,44c	-0,35c
8	3,95c	4,12c	-0,17c
9	4,33c	4,85c	-0,52c
10	4,01c	4,53c	-0,52c
11	4,39c	5,07c	-0,68c
12	4,38c	4,18c	0,20c
13	4,30c	4,33c	-0,03c
14	4,90c	5,47c	-0,57c
15	4,43c	4,50c	-0,07c
	Ø 4,17c	Ø 4,38c	Ø -0,21c

Tab. 36: Datenübersicht zur Lage des zweiten Peaks im Vergleich der manuellen und der automatisierten Messung für Tribukait C-Fälle

Die Tests auf Normalverteilung zeigten, dass die Daten normalverteilt sind und somit die parametrische Statistik anwendbar war. In Tabelle 37 sind die entsprechenden Daten dargestellt.

Aus den Daten und der Wertetabelle kann man erkennen, dass die Messwerte des Peak 2 der Diagnosegruppe Tribukait C eine noch recht gute Übereinstimmung zeigen. Auch der Bland-Altman-Plot (Abb. 32) weist alle Wertepaare im Toleranzbereich aus, außerdem lässt sich ablesen, dass die gemessenen Schwankungen offenbar unabhängig von der Lage der Messwerte waren. Der t-Test bzgl. der Lage des Mittelwertes der beiden einzelnen Messwerte im Vergleich zur Null-Linie weist allerdings mit $p = 0,019$ darauf, dass der Unterschied von im Durchschnitt $-0,21c$ statistisch signifikant ist.

Spearman's Korrelationskoeffizient		t-Test der Stichproben		Konfidenzintervall des Bias		Konfidenzintervall untere Grenze der Übereinstimmung (95% Konfidenz)		Konfidenzintervall obere Grenze der Übereinstimmung (95% Konfidenz)	
<i>Spearman's Rho</i>	<i>p</i> -wert	t	<i>p</i> -wert	untere	obere	untere	obere	untere	obere
0,271	0,328	2,659	0,019	0,0414	0,3866	-0,6922	-0,1018	0,5298	1,1202

Tab. 37: Teststatistik zum Vergleich der Lage des Peak 2 im Vergleich automatischer und manueller Messungen für Tribukait C-Fälle.

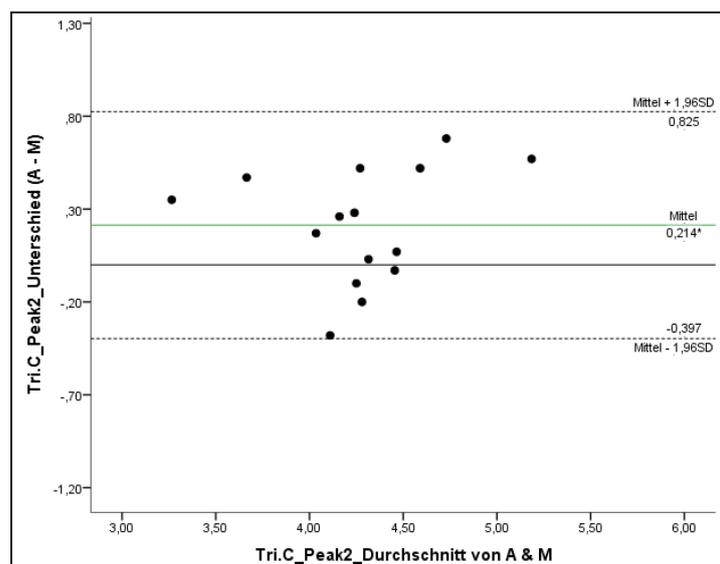


Abb. 32: Bland-Altman-Plot für Peak 2 im Vergleich der manuellen Messung (M) und der ersten automatisierten Messung (A) der Tribukait C-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

Vergleich zwischen den beiden automatisierten Messungen

Die Daten für den Vergleich der beiden automatisierten Messungen sind in Tabelle 38 wiedergegeben. Bei Fall 13 ließ sich die zweite Messung nicht auswerten, so dass der Fall hier nicht berücksichtigt werden konnte. Insgesamt waren sich die Messwerte weitgehend gleich; die durchschnittlichen Werte lagen bei 4,38c bzw. bei 4,28c und unterschieden sich somit nur um 0,10c. Die größte Differenz zwischen zwei Messungen betrug allerdings doch 0,71c. Sieben Messwerte der ersten und fünf Messwerte der zweiten automatisierten Messung zeigten sich außerhalb des Toleranzbereiches von 4c +/- 10% und führten zur Einordnung als aneuploide DNA-Stammlinie. Bei sieben Fällen lag der erste Messwert höher, bei sechsen der zweite, in einem Fall waren beide Messwerte exakt gleich; insgesamt ergab sich hier kein auf Unterschiede zwischen den Messreihen weisender Trend.

Tribukait C			
Fall	Automatisierte Messung 1	Automatisierte Messung 2	Differenz
1	4,38c	4,27c	0,11c
2	3,90c	3,59c	0,31c
3	4,44c	4,30c	0,14c
4	3,92c	3,98c	-0,06c
5	4,20c	4,20c	0,00c
6	4,29c	4,31c	-0,02c
7	3,44c	3,51c	-0,07c
8	4,12c	4,29c	-0,17c
9	4,85c	4,61c	0,24c
10	4,53c	4,54c	-0,01c
11	5,07c	4,96c	0,11c
12	4,18c	4,40c	-0,22c
13	4,33c	-	-
14	5,47c	4,76c	0,71c
15	4,50c	4,22c	0,28c
	Ø 4,38c	Ø 4,28	Ø 0,10c

Tab. 38: Datenübersicht zur Lage des zweiten Peaks im Vergleich der beiden automatisierten Messungen für die Tribukait C-Fälle. Im Mittelwert der automatisierten Messung 1 ist der Wert von Fall 13 nicht berücksichtigt.

Die Tests auf Normalverteilung zeigten, dass auch an diesen Daten eine parametrische Statistik angewendet werden konnte. In Tabelle 39 sind die entsprechenden Daten wiedergegeben.

Spearman's Korrelationskoeffizient		t-Test der Stichproben (Bias null)		Konfidenzintervall des Bias		Konfidenzintervall untere Grenze der Übereinstimmung (95% Konfidenz)		Konfidenzintervall obere Grenze der Übereinstimmung (95% Konfidenz)	
<i>Spearman's Rho</i>	<i>p</i> -wert	t	<i>p</i> -wert	untere	obere	untere	obere	untere	obere
0,497	0,070	1,156	0,153	-0,0410	0,2338	-0,6049	-0,1351	0,3279	0,7977

Tab. 39: Teststatistik zum Vergleich der Lage des Peak 2 im Vergleich der beiden automatisierten Messungen für Tribukait C-Fälle.

Aus den Daten und der Wertetabelle ergibt sich, dass die Messwerte des Peak 2 der Diagnosegruppe Tribukait C weitestgehend als übereinstimmend angesehen werden können. Auch der Bland-Altman-Plot (Abb. 33) zeigt im Wesentlichen keine auffälligen Schwankungen der Messwerte; ein Wertepaar liegt allerdings außerhalb des Vertrauensbereiches. Eine Abhängigkeit der gemessenen Unterschiede von der Höhe der Messwerte ist ebenfalls nicht abzuleiten.

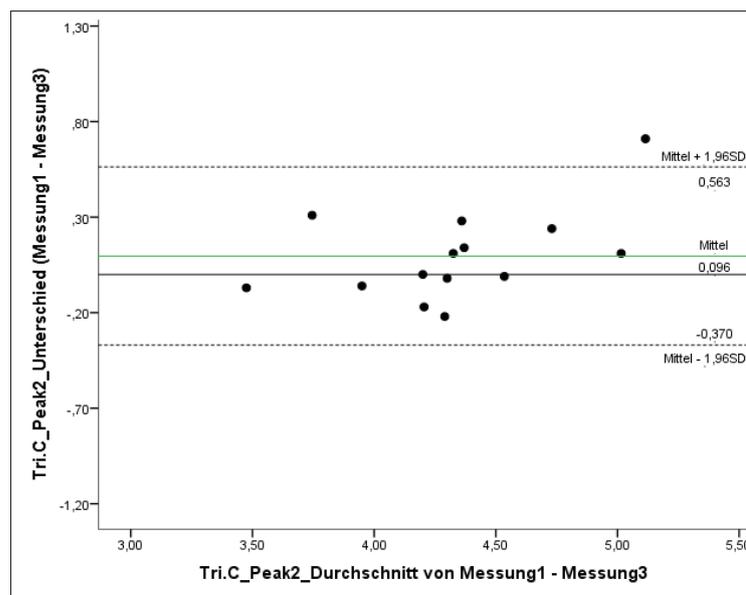


Abb. 33: Bland-Altman-Plot für Peak 2 im Vergleich der beiden automatisierten Messungen der Tribukait C-Fälle. x-Achse: Mittelwert der beiden Messwerte, y-Achse: Differenz der beiden Messwerte.

4.4 Klinische Datenauswertung

4.4.1 Altersverteilung der Patienten

Von den 27 Fällen waren in der manuellen DNA-zytometrischen Messung 12 als Typ B nach Tribukait („Tribukait B“) und 15 als Typ C nach Tribukait („Tribukait C“) eingestuft worden. Die in Abb. 34 dargestellte Altersverteilung lässt für beide Typen eine erwartete Häufung bei Patienten in der Altersklasse von 61-70 Jahren erkennen; statistisch signifikante Unterschiede bestanden nicht (Pearson-Chi²-Test: $p > 0,05$).

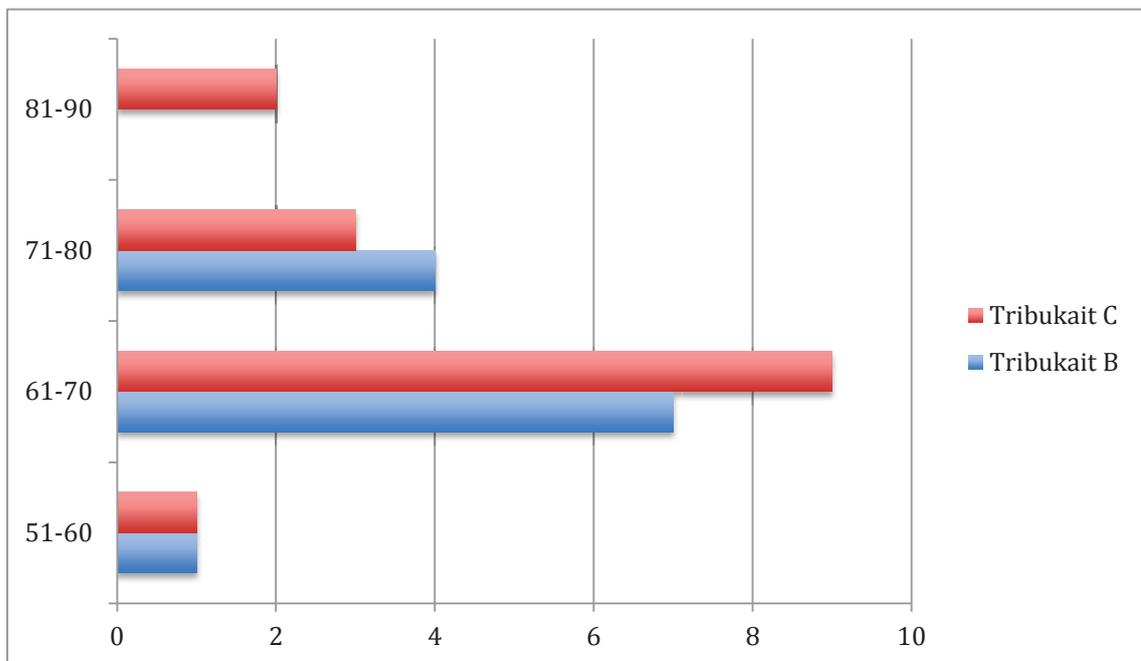


Abb. 34: Altersverteilung der Patienten in 10-Jahres-Gruppen, geordnet nach den Tribukait-Typen

4.4.2 Gegenüberstellung der Ergebnisse von manueller und automatisierter DNA-Bildzytometrie

Da die beiden automatisierten DNA-bildzytometrischen Messungen in den 26 Fällen, in denen zwei Messungen erfolgen konnten, jeweils zu einem gleichlautenden Ergebnis bezüglich des Tribukait-Typs kamen, können die

klinischen Ergebnisse im Folgenden vereinfacht als Gegenüberstellung von manueller Messung und automatisierter Messung dargestellt werden; eine Fallunterscheidung nach den automatisierten Messungen 1 und 2 ist nicht notwendig. Alle 53 automatisierten Messungen ergaben einen Typ B oder Typ C nach Tribukait, Einstufungen als Typ A oder Typ D kamen nicht vor.

Die Ergebnisse der manuellen und der automatisierten DNA-bildzytometrischen Messungen wurden in einer Vierfeldertafel gegenübergestellt (Tabelle 40). Neun der 12 Fälle mit manuell ermitteltem Typ B wurden automatisiert gleichartig eingeordnet (75%); beim Typ C wurden 13 von 15 Fällen gleich eingestuft (86,7%). Insgesamt lagen somit in 22 Fällen konkordante Ergebnisse vor (81,5%); der Kappa-Wert lag bei 0,622, was einer „guten Übereinstimmung“ entspricht.

Automatisierte DNA-Bildzytometrie	Manuelle DNA-Bildzytometrie		Total
	B	C	
B	9	2	11
C	3	13	16
Total	12	15	27

Tab. 40: Übereinstimmungsmatrix von manueller und automatisierter DNA-bildzytometrischer Messung bei Klassifikation von 27 Prostatakarzinomen als Typ B oder Typ C nach Tribukait

In fünf der 27 Fälle lag keine Übereinstimmung vor. Zwei Fälle wurden in der automatisierten Messung als Tribukait C statt in der manuellen Messung als Tribukait B, also höher eingestuft, in drei Fällen war es umgekehrt.

4.4.3 Diskrepanz beurteilte Fälle

Die zur Diagnose eines Typ B bzw. Typ C führenden DNA-zytometrischen Kenndaten der fünf diskrepanten Fälle sind in Tabelle 41 dargestellt.

Fall / Mess.		Tribukait		Peak 1		Peak 2		Peak 3		C-Max	
		Auto.	Man.	Auto.	Man.	Auto.	Man.	Auto.	Man.	Auto.	Man.
1	1	B	C	2,17	2,04	3,90	3,43	-	-	4,47	5,30
1	2	B		2,10		3,59				4,78	
2	1	B	C	2,05	2,20	4,20	4,30	-	-	8,38	10,94
2	2	B		2,00		4,20				7,25	
3	1	C	B	2,00	2,02	4,70	4,03	-	-	5,67	6,11
3	2	C		2,00		4,69				5,22	
4	1	C	B	2,07	2,05	4,38	4,27	5,09	-	6,33	7,23
4	2	C		2,05		4,14		4,63		5,23	
5	1	C	B	2,00	1,99	4,14	3,98	4,91	-	7,41	8,63
5	2	C		1,97		4,06		4,80		7,62	

Tab. 41: DNA-zytometrische Kenndaten der beiden automatisierten und der singulären manuellen Messung von fünf diagnostisch diskrepanten Fällen

Im Folgenden sollen auf die Unterschiede in den DNA-Messwertverteilungen eingegangen werden und die diagnostischen Interpretationen skizziert werden.

Fall 1:

Die DNA-Histogramme zeigen ein ähnliches Grundmuster mit einer diploiden DNA-Stammlinie, einem zweiten Peak im Bereich im tetraploiden Bereich um 4c und Einzelzellen > 4,4c: Diagnostisch ausschlaggebend ist die Interpretation des zweiten Peaks. Er liegt bei der manuellen Messung außerhalb von 4c +/- 10% (= Typ C), bei den beiden automatisierten Messungen innerhalb bzw. auf der Grenze (= Typ B).

Fall 2:

Die DNA-Histogramme weisen auch hier jeweils eine diploide DNA-Stammlinie auf, begleitet von einem zweiten Peak im Bereich im tetraploiden Bereich um 4c und von Einzelzellen > 4,4c: Bei den beiden automatisierten Messungen reichen die Einzelzellen bis 7,25c bzw. 8,38c, bleiben also unterhalb der zu einer Interpretation als Typ C liegenden Schwelle von 9c, während manuell eine Zelle mit einem DNA-Gehalt von 10,94c ermittelt wurde. Daraufhin wurden die Bildgalerien verglichen, die aber nur bedingt hilfreich waren, Zellkerne anhand der Morphologie vergleichend zwischen verschiedenen Messungen zu identifizieren und daher keine echte Aufklärung der Situation ermöglichten.

Fall 3:

Alle drei Messungen zeigen eine diploide DNA-Stammlinie sowie einen zweiten Peak, der bei der manuellen Messung mit 4,03c im tetraploiden Bereich lag (Typ B), bei den automatisierten Messungen hingegen mit 4,70c und 4,69c im aneuploiden Bereich von außerhalb von $4c \pm 10\%$ (Typ C). Zellen mit Einzelwerten $> 9c$ kamen bei allen drei Messungen nicht vor.

Fall 4:

Die DNA-Histogramme zeigen ein ähnliches Grundmuster mit einer diploiden DNA-Stammlinie und einem zweiten Peak im Bereich im tetraploiden Bereich um $4c$. Hinzu kommt aber bei beiden automatisierten Messungen ein dritter Peak außerhalb des $4c$ -Bereiches, nämlich bei 5,09c bzw. bei 4,63c. Zellen mit Einzelwerten von $> 9c$ wurden nicht nachgewiesen. Der dritte Peak außerhalb des $4c$ -Bereiches führt zur Interpretation der automatisierten Messungen als Typ C, während die manuelle als Typ B ausgewertet wird.

Fall 5:

Die Unterschiede in den DNA-Histogrammen gleichen denjenigen bei Fall 4. Es traten auch in diesem Fall bei den automatisierten Messungen dritte Peaks außerhalb des $4c$ -Bereiches auf, nämlich bei 5,09c bzw. bei 4,63c, vor, die zur Einordnung als Typ C führten. Bei der manuellen Messung lag ein solcher dritter Peak nicht vor (= Typ B).

5 Diskussion

5.1 Das Prostatakarzinom - ein „Sonderfall“ in der klinischen Tumormedizin

Das Prostatakarzinom nimmt innerhalb der Karzinome in mehrererlei Hinsicht eine Sonderstellung ein.

Seit langem ist anerkannt, dass es sich im Wesentlichen um einen Tumor älterer Männer handelt. In einem aktuellen Review, in dem 29 zwischen 1948 und 2013 veröffentlichte, jeweils mindestens 100 Fälle aufweisende Obduktionsstatistiken aus Europa, Asien und Amerika zusammengestellt worden waren, wurde bei insgesamt 8583 Fällen und in allen Studien eine altersabhängig steigende, sich alle 14 Jahre verdoppelnde adjustierte Prävalenz von Prostatakarzinomen angegeben (Bell et al. 2015): Bei Männern von < 30 Jahren betrug sie 5%, bei Männern ab 80 Jahren lag sie im Mittel bei 59%. Zu Ende gedacht führt diese Steigerung zu der Überlegung, ob nicht letztendlich jeder Mann bei „unendlichem“ Leben an einem Prostatakarzinom erkranken würde.

Ebenfalls war bereits bekannt, dass Prostatakarzinome letztendlich über eine gute Prognose verfügen. Die aktuellen Zahlen des Robert-Koch-Institutes geben für Deutschland, bezogen auf 2011/2012, bei einem medianen Erkrankungsalter von 71 Jahren absolute 5-Jahres-Überlebensraten und 10-Jahres-Überlebensraten von 78% bzw. 60% an; die entsprechenden relativen, also auf die allgemeine Sterbetafel von 71-Jährigen bezogenen Werte liegen bei 93% bzw. 91%. Berücksichtigt sind dabei alle Sterbefälle, nicht etwa nur die durch ein Prostatakarzinom bedingten Fälle (RKI 2015). Diese Zahlen unterstreichen die im Allgemeinen sehr gute Prognose von Prostatakarzinomen; nur für maligne Hodentumoren ergibt sich mit 96% bzw. 95% ein besseres relatives 5- bzw. 10-Jahres-Überleben.

Die statistisch sehr gute Prognose ist allerdings kritisch im Gegenlicht sich verändernder epidemiologischer Zahlen zu sehen. In einer früheren Auflage der Krebsdatenübersicht vom Robert-Koch-Institut, bezogen auf die Zahlen von

2005/2006, wird die „Entwicklung von Inzidenz und Mortalität“ wie folgt beschrieben (RKI 2010):

„Die Zahl jährlich neu aufgetretener Krebserkrankungen der Prostata in Deutschland ist seit 1980 um 200 %, die altersstandardisierte Erkrankungsrate um 110 % angestiegen. Dieser Anstieg kann größtenteils auf den Einsatz neuer Methoden in der Diagnostik, z. B. der Bestimmung des prostataspezifischen Antigens (PSA), zurückgeführt werden. Die frühzeitigere, hinsichtlich der Entwicklung der Krebserkrankung, aber auch in Bezug auf das Alter der Patienten gestellte Diagnose, führt zu erheblich höheren Erkrankungsraten im Altersbereich zwischen 50 und 69 Jahren bei gleichzeitig abnehmenden Erkrankungsrate über 75-Jähriger und einem abnehmenden mittleren Erkrankungsalter von 73 Jahren 1980 auf 69 Jahre 2006. Die Zahl der Sterbefälle hat in Folge der demografischen Entwicklung, d. h. der Zunahme älterer Männer in der Bevölkerung, seit 1980 zwar um 30 % zugenommen, die altersstandardisierte Sterberate ist aber um 20 % gefallen.“

Der dramatische Anstieg von diagnostizierten Prostatakarzinomen lässt sich auch aus Abb. 35 ablesen.

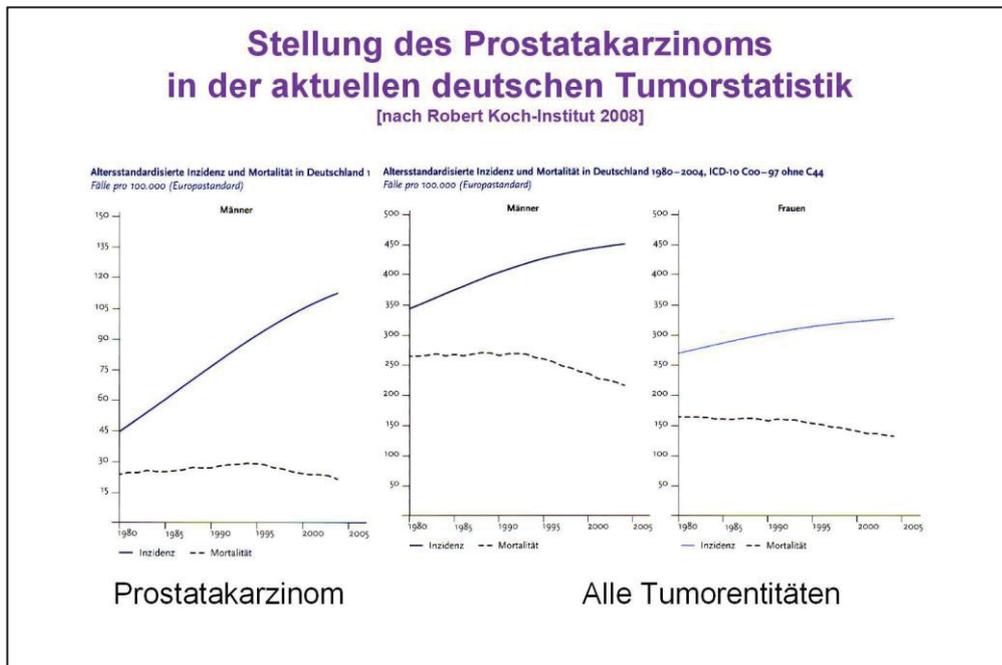


Abb. 35: Altersstandardisierte Inzidenz von Prostatakarzinomen sowie allen Malignomen bei Männern und Frauen in Deutschland zwischen 1980 und 2005 (Biesterfeld 2014 nach RKI-Daten (RKI 2012)); Details siehe folgenden Text.

Sie illustriert seinen Anstieg von ca. 45 Fällen / 100.000 Männer im Jahr 1980 auf ca. 115 Fälle im Jahr 2005. Setzt man diese Zahlen in das Verhältnis zum Anstieg aller malignen Tumoren bei Männern im gleichen Zeitraum (von ca. 350 auf ca. 450), so gehen ca. 70% des Anstiegs an Tumorfällen auf das Konto neu entdeckter Prostatakarzinome.

Die stark gestiegene Inzidenz wurde frühzeitig bereits als Effekt neu entstandener diagnostischer Möglichkeiten, insbesondere der standardisierten Prostatastanzbiopsie bei erhöhtem PSA-Wert, gedeutet und nicht als Indikator für eine wirklich vermehrte Zahl von Neuerkrankungen gesehen. Entsprechend ging man davon aus, dass bei vielen Patienten Prostatakarzinome diagnostiziert wurden, die ihnen vielfach über ihre gesamte Lebenszeit keine klinischen Probleme gemacht hätten, wenn sie unentdeckt geblieben wären. Im Grunde bedeutete die höhere Detektionsrate der Karzinome eine Annäherung der Krankheitszahlen an die oben erwähnten Tumorraten in den Obduktionsstatistiken.

Dass es Patienten gab, die ein Prostatakarzinom sehr lange Zeit überlebten und nicht „am“, sondern „mit dem“ Prostatakarzinom verstarben, war ebenso bekannt wie die Tatsache, dass andere Patienten schnell nach Metastasierung an diesem Tumor verstarben. Hieraus resultierten bereits in den 70er Jahren die Wortschöpfungen des „Haustierkrebses“ und des „Raubtierkrebses“, die von Hackethal eingeführt bzw. auf das Prostatakarzinom angewandt wurden. Gerade das Wissen um die schlechten schnellen Verläufe der „Raubtierkrebse“ führte dazu, dass in der Urologie eine Zeitlang der hohen Zahl von Neuerkrankungen mit erheblichen therapeutischen Anstrengungen begegnet wurde, obwohl klar war, dass man damit bei der guten Gesamtprognose vielen Patienten eine Übertherapie anbot.

Was es allerdings ebenfalls bereits lange gab, war die Strategie des heute „Watchful Waiting“² genannten therapeutischen Abwartens bei Patienten, die

²Eine Definition des *Watchful Waiting* ergibt sich aus Empfehlung 5.11 der S3-Leitlinie: „Langfristige Beobachtung und sekundäre symptomorientierte Therapie („Watchful Waiting“) statt kurativer Behandlung soll bei Patienten erörtert werden, die eine mutmaßliche Lebenserwartung unter zehn Jahren haben.“

aus anderen Gründen über eine nur geringe Lebenserwartung verfügten, weil sie beispielsweise sehr alt waren oder an anderen, die Prognose einschränkenden bzw. einer Therapie hinderlichen chronischen Erkrankungen litten. Rationale für eine abwartende Haltung ist dabei, dass das Prostatakarzinom des Patienten oft nur langsam voranschreitet und die Komorbiditäten im Grunde den klinischen Gesamtverlauf in einem höheren Maße bestimmen (Adolfsson 2008, Mottet et al. 2017).

Für die anderen Patienten wurden hormonelle Verfahren, strahlentherapeutische Verfahren bzw. die radikale Prostatektomie vorgesehen. Gerade die Zahl letzterer Eingriffe nahm in den 90er Jahren nach Erfahrungen der Pathologen im Krankengut der Urologischen Kliniken deutlich zu; auch das Ausmaß ihrer Aufarbeitung veränderte sich von orientierenden Schnittführungen hin zur kompletten Einbettung, welche heute Standard ist. In der Empfehlung 4.35 der S3-Leitlinie wird dazu unverändert seit 2009 explizit u.a. ausgeführt, "das Prostatektomiepräparat (solle) in der transversalen Ebene in 3-5 mm dicke Scheiben lamelliert und komplett eingebettet werden (Großflächenschnitte) oder orientiert in normalen Histologiekapseln" (Leitlinie Onkologie 2016). Man verspricht sich davon eine sicherere Beurteilung des Tumorstadiums (T2 vs. T3a oder T3b) und eine besser begründete Aussage darüber, ob der Resektionsrand im Gesunden ist (R0 vs. R1) (Desai et al. 2002).

Sowohl die Hormontherapie, als auch die Strahlentherapie und die radikale Prostatektomie bringen aber zu diskutierende Nebenwirkungen mit sich. Exemplarisch soll hier auf diejenigen der radikalen Prostatektomie als dem operativen Standardverfahren eingegangen werden. Erwähnt werden im Kapitel 5.3 der S3-Leitlinie („Lokale Therapie des lokal begrenzten Prostatakarzinoms“) therapierrelevante Blutungen, rektale Verletzungen, tiefe Beinvenenthrombosen und Lungenembolien, Lymphozelen, Urin-Leckagen und Fisteln, geringgradige bzw. hochgradige Belastungsinkontinenz, erektile Dysfunktion, Blasenhalss- bzw. Ureterobstruktion und Urethrastrikturen. Die im Kapitel 13.3 („Vergleichende Zusammenstellung des Nebenwirkungsprofils der Verfahren zur Prostatakarzinombehandlung“) tabellarisch zusammengefassten prozentualen Werte

aus verschiedenen Studien und Leitlinien für alle kurativen Therapieansätze, insbesondere auch unter Einbeziehung der strahlentherapeutischen Verfahren, sind durchaus bedenkenswert, auch wenn sie stark von Studie zu Studie schwanken. In einem neueren systematischen Review zu den Nebenwirkungen der radikalen Prostatektomie wurden beispielsweise Werte für erektile Dysfunktion zwischen <5% und 86% herausgearbeitet (Wilt et al. 2008), also mit im Grunde unverständliche Spannweite der Daten. Im Durchschnitt der Studien dürften sie über alle Verfahren bei etwa 30%-50% liegen. Für eine persistierende Harninkontinenz werden bei der radikalen Prostatektomie Werte von 5%-10% genannt (Dutch Urological Association 2007), für strahlentherapeutische Verfahren von bis zu 20% (Heidenreich et al. 2008).

Setzt man nun die nicht selten kleinen und auf das Organ beschränkten Tumoren mit ggfls. blandem Verlauf in Beziehung zum Nebenwirkungsprofil der Therapien, so wird verständlich, dass nach Alternativen zu einer definitiven Tumorthherapie gesucht wurde. Hieraus leitet sich unter anderem auch das Konzept der „Active Surveillance“ ab.

5.2 Grundzüge der „Active Surveillance“-Strategie

Folgt man den Vorgaben der S3-Leitlinie (Leitlinienprogramm Onkologie 2016/, so bedeutet, den Weg der „*Active Surveillance*“ (AS) einzuschlagen,

- zunächst auf eine Therapie zu verzichten, und zwar unter der Annahme, dass zum Zeitpunkt der Diagnosestellung ein Prostatakarzinom mit nur geringem Malignitätspotential und entsprechend geringem Progressionsrisiko vorliegt, sowie
- durch engmaschige, Leitlinien-gerechte Nachsorge dafür Sorge zu tragen, dass eine eventuelle Zunahme des Malignitätspotentials so zeitig entdeckt wird, dass eine kurative Tumorthherapie mit dem Ziel einer definitiven Heilung noch rechtzeitig eingeleitet werden kann.

Die Abgrenzung einer *Active Surveillance*-Strategie gegenüber einer klassischen Therapie liegt also darin, bei Diagnosestellung des

Prostatakarzinoms zunächst auf eine Therapie zu verzichten. Gegenüber dem *Watchful Waiting* besteht die Grenzziehung in dem kurativen Ansatz, der für *Active Surveillance* gegeben ist, für *Watchful Waiting* aber nicht.

Das Prostatakarzinom ist somit der einzige solide Organtumor, bei dem nicht mehr in jedem Fall eine zeitnahe Tumorthherapie erfolgen muss, anders als bei anderen epidemiologisch wichtigen Tumorentitäten wie beispielsweise Karzinomen der Lunge, der Mamma oder des Kolorektums bzw. malignen Melanomen oder Non-Hodgkin-Lymphomen. Diese Option, die die Patienten nicht gewohnt sind, schafft bei ihnen, aber auch bei den sonst stets auf Behandlung eingestellten klinischen Ärzten eine gewisse Unsicherheit über die Bedingungen und Möglichkeiten der AS-Strategie.

In der aktuellen S3-Leitlinie, die von wenigen Modifikationen abgesehen auch bereits bei Diagnosestellung der in unserer Studie analysierten Fälle galt, wird dazu im Kapitel 5 „Therapie des nichtmetastasierten Prostatakarzinoms“ in den Empfehlungen 5.7³ bis 5.10 wie folgt ausgeführt (Leitlinien Onkologie 2016):

- Empfehlung 5.7: „Patienten mit einem lokal begrenzten Prostatakarzinom, für die eine kurativ intendierte Behandlung in Frage kommt, sollen nicht nur über Behandlungsverfahren wie radikale Prostatektomie und perkutane Strahlentherapie sowie Brachytherapie, sondern auch über Aktive Überwachung (*Active Surveillance*) informiert werden.“

- Empfehlung 5.8: „a. Voraussetzung für die Wahl einer Strategie der Aktiven Überwachung sollen folgende Parameter sein:
 - PSA-Wert ≤ 10 ng/ml;
 - Gleason-Score ≤ 6 ;
 - cT1 und cT2a;
 - Tumor in ≤ 2 Stenzen bei leitliniengerechter Entnahme von 10-12 Stenzen
 - ≤ 50 % Tumor pro Stanze.

- b. Bei Gleason 3+4 (7a) sollte die Aktive Überwachung im Rahmen von Studien geprüft werden.“

³Die Nomenklatur und die Zählweise der Empfehlungen folgen der Nummerierung in der aktuellen S3-Leitlinie.

- Empfehlung 5.9: „a. Der Tumor soll in den ersten beiden Jahren durch PSA-Bestimmung und DRU alle drei Monate kontrolliert werden. Bleibt der PSA-Wert stabil, ist 6-monatlich zu untersuchen. Eine Rebiopsie soll erfolgen.
b. Sie sollte nach 6 Monaten durchgeführt werden.
c. Biopsien sollen danach in den ersten drei Jahren alle zwölf bis 18 Monate vorgenommen werden, später bei stabilem Befund alle drei Jahre.“
- Empfehlung 5.10: „Wenn die Einschlusskriterien in einem Kriterium nicht mehr erfüllt sind, oder sich die PSA-Verdopplungszeit auf weniger als drei Jahre verkürzt, soll zu einer Beendigung der Aktiven Überwachung geraten werden.“

Damit liegen sowohl Empfehlungen für die Entscheidung für oder gegen eine *Active Surveillance*-Strategie als auch für ihre Weiterführung bzw. ihren Abbruch im Laufe der Zeit vor, die eine Leitlinien-bezogene Vorgehensweise möglich machen sollten.

Nicht mehr in der S3-Leitlinie enthalten ist die in der Vorgängerversion von 2011 verankerte Empfehlung (damals 5.7), dass „Bestandteil der Aufklärung (...) neben den Vorteilen auch die fehlende Verfügbarkeit von Langzeitdaten und die Unsicherheit der Ergebnisse aufgrund kleiner Studienpopulationen sein“ solle. Dem ist offenbar durch Studien inzwischen Genüge getan: Vier prospektive Studien mit jeweils mehr als 90 Patienten, die in der S3-Leitlinie in den sogenannten „Hintergrundinformationen“ explizit dargestellt werden sind, zeigen zwar ein heterogenes Bild von Einschlusskriterien, dokumentieren aber die mit einer *Active Surveillance*-Strategie zu erzielenden Ergebnisse (Al Otaibi et al. 2008, Khatami et al. 2007, Klotz et al. 2010, Shapley et al. 2009). Dargestellt wurden bei mittleren Beobachtungszeiträumen zwischen 5 1/4 Jahren und 6 1/2 Jahren die sehr ähnlichen Tumor-spezifischen Überlebensraten (97,6% - 100%), das recht heterogene Gesamtüberleben (78,6% - 97,8%) sowie auch Details zu Rezidiven, Metastasenbildung und allgemein zur Tumorprogression. In 30% - 49% der Fälle wurde im Laufe des jeweiligen Beobachtungszeitraumes auf eine definitive Therapie umgestellt; die übrigen Patienten konnten über eine längere Zeit begründet therapiefrei bleiben.

5.3 DNA-Bildzytometrie als ergänzendes Auswahlkriterium für die Wahl einer „Active Surveillance“-Strategie

Die Kriterien, nach denen ein Patient für eine *Active Surveillance*-Strategie geeignet sein soll, sind nicht unwidersprochen geblieben und verfügen über Schwachstellen, die nicht immer beachtet werden.

Zunächst ist zu sagen, dass bei der Definition sehr konservativ vorgegangen wurde und sich die gewählten Kriterien weitgehend an einer einzigen Übersichtsarbeit orientieren (Klotz 2005). Andere Ansätze wurden nicht oder nur sehr zurückhaltend aufgenommen. So wäre beispielsweise nach der NICE-Leitlinie von 2008 auch vertretbar, einen PSA-Wert von 15 ng/ml für *Active Surveillance* zu akzeptieren (NICE 2008). Hinter der zurückhaltenden Strategie der Leitlinienkommissionen steht sicher der verständliche Wunsch des *nil nocere*, also des Prinzips, Schaden vom Patienten abzuwenden. Dennoch bleibt die Frage, ob minimales Überschreiten von Kriterien, also ein PSA-Wert von 11 ng/ml oder ein Gleason-Grad 7a oder eine dritte positive Stanze mit minimalem Tumor, wirklich als Ausschlusskriterium zu bewerten ist.

Wichtiger ist vielleicht, anzuführen, dass auch die jetzt gewählten Kriterien bereits bei ihrer Erfüllung Probleme mit sich bringen.

Dieses gilt zum einen für die Beschränkung des Tumorstadiums auf höchstens T2a, also auf den Befall von < 50% des Volumens eines der beiden Prostatalappen. Fasst man Erfahrungen aus Prostatektomien auch kleiner Prostatakazinome zusammen, so lässt sich sagen, dass kaum einmal eine Beschränkung auf einen Lappen vorliegt, sondern in der großen Mehrzahl der Fälle bei kleinherdig-diskontinuierlichem Wachstumsmuster ein Stadium pT2c besteht. Der Sinn der Beschränkung auf eine T2a-Situation erschließt sich somit nicht klar, und warum Patienten, die sonst alle Kriterien für eine *Active Surveillance*-Strategie erfüllen, deren beiden Tumor-positiven Stanzen aber auf beide Lappen verteilt sind, hiervon ausgeschlossen werden sollen, auch nicht.

Von besonderer Bedeutung sind aber die Probleme mit dem Gleason-Grading, auf dessen eingeschränkte Reproduzierbarkeit auch in der S3-Leitlinie explizit hingewiesen wird. Schon der ursprüngliche Gleason-Score waren die Werte zwischen zwei Untersuchern mit 36% - 81% für genaue Übereinstimmung und 69% - 86% für eine Übereinstimmung von +/- einem Scorewert eher mäßig (Allsbrook et al. 2001, Rubin et al. 2004, Sakr et al. 2000). Für den modifizierten Gleason-Score (Epstein et al. 2005) ist eine Verbesserung dieser Werte nicht allgemein anerkannt (Burchardt et al. 2008, Kuroiwa et al. 2010, Netto et al. 2011, Veloso et al. 2007), auch wenn in manchen Studien eine Verbesserung der Reproduzierbarkeit gesehen wurde (Helpap und Egevad 2007, Melia et al. 2006). Auch die Frage, ob die durch die Modifikation angestrebte höhere Korrelation zwischen den Ergebnissen in Stanzbiopsien und späteren Prostat-ektomien realisiert werden konnte, ist bislang bei kontroversen Ergebnissen nicht klar zu beantworten (Helpap und Egevad 2006, Egevad et al. 2012, Uemura et al. 2009, Kuroiwa et al. 2010, Veloso et al. 2007, Zareba et al. 2009).

Das Schlagwort „3 Pathologen - 2 Meinungen“ hat somit leider für das Grading des Prostatakarzinoms nichts an Brisanz und Bedeutung verloren. Urologen ist dieses Themenfeld leider sehr fremd, wie eine neuere Untersuchung einer deutschen urologischen Arbeitsgruppe an Befundberichten aus der Pathologie ergab (Damjanoski et al. 2013). Inwieweit eine standardmäßige Referenzpathologie oder zumindest eine Zweitbegutachtung auf fachärztlicher Basis zu besseren Ergebnissen führen, ist nicht bekannt. Klar ist aber, dass auch unter Pathologen die Bedeutung mancher Gleason-Grade in seiner biologischen Wertigkeit umstritten ist, wie ihre recht willkürliche Übersetzung in das in der Pathologie übliche System als G1, G2 oder G3 belegt (Biesterfeld 2015).

Die Modifikation des Gleason-Gradings birgt noch ein weiteres Problem. Die Definition der Kriterien für eine *Active Surveillance*-Strategie basieren noch auf Untersuchungen mit dem früheren Gleason-Grading. Wenn nun durch die Modifikation der Definition häufiger Fälle als 7a statt bisher als 6 eingestuft werden, ohne dass die Kriterien angepasst werden, dann wird Patienten eine

Therapieoption ggfls. weggenommen, obwohl sie die Voraussetzungen lediglich nicht formal, inhaltlich aber sehr wohl erfüllen.

Hier nun kommt die DNA-Bildzytometrie ins Spiel. Auch wenn diese nur an wenigen Zentren durchgeführt wird, ist sie doch eine anerkannte Methode, die anders als manche andere Anwendung seit vielen Jahren in den Leistungskatalog der gesetzlichen Krankenkassen aufgenommen ist und somit jedem Patienten prinzipiell zur Verfügung steht. Von ihr erhofft man sich gerade auch seitens der Selbsthilfeverbände der Prostatakarzinom-Patienten, insbesondere auch beim Dachverband „Bundesverband Prostatakrebs Selbsthilfe“ (BPS), ergänzende Informationen darüber, ob ein individueller Patient für eine *Active Surveillance*-Strategie geeignet ist oder nicht. Damit eine solche zusätzliche Untersuchung einen Sinn macht, muss sie verschiedene Anforderungen erfüllen. Dazu gehören in erster Linie eine akzeptable Validität und eine hohe Reproduzierbarkeit von Messergebnissen verschiedener Untersucher und ferner insbesondere die morphologische Nachvollziehbarkeit dessen, was methodisch und inhaltlich umgesetzt wird, also die Möglichkeit, intellektuell das biologische Konzept hinter der Methode zu erfassen.

Gerade in Letzterem liegt ein großer Vorteil der DNA-Bildzytometrie. Seit der Standardisierung der Anforderungen an Labormethoden und Messmethoden (Böcking et al. 1995, Giroud et al. 1998, Haroske et al. 1998, Haroske et al. 2001) ist auch allgemein akzeptiert, was diese Anwendung leisten kann und wo ihre Grenzen liegen. Das biologische Konzept, über die Bestimmung des mittleren DNA-Gehaltes der G₀/G₁-Phase der Tumorzellpopulation auf eine pathologische numerische chromosomale Aberration zu schließen, dafür aber nicht an den Chromosomen selber untersuchen zu müssen, ist im Grunde einleuchtend und verständlich und inzwischen auch unter den die Methode betreibenden Gruppen anerkannt.

In der S3-Leitlinie wird der Stellenwert der DNA-Bildzytometrie in der Wahl einer *Active Surveillance*-Strategie zurückhaltend gesehen. In den Erläuterungen zu Empfehlung 5.8 heißt es:

„Auch für die DNA-Zytometrie stehen nach Ansicht der Autoren nicht genügend Daten zur Verfügung, um den Routine-Einsatz als prädiktiven Marker für eine aktive Therapie zu begründen.“

Diese Einschätzung ist aus Sicht einer aktiven DNA-bildzytometrischen Arbeitsgruppe nicht zu beanstanden. Gerade wenn es Kriterien für eine *Active Surveillance*-Strategie gibt, wäre sowohl bei Patienten, die diese klar erfüllen, als auch bei Patienten, die diese klar nicht erfüllen, eine solche ergänzende Untersuchung in der Tat nicht vonnöten, vorausgesetzt, dass sich die Tumordifferenzierung nach dem Gleason-Grading im Konsens mehrerer Fachärzte erbringen ließe. Entsprechend geht es um diese Patienten eigentlich nicht, sondern um diejenigen, die „auf der Kippe“ stehen. Als Beispiele könnten Konstellationen dienen, wie sie oben bereits zu Beginn des Kapitels 5.3 beim Thema des minimalen Überschreitens der Kriterien für eine *Active Surveillance*-Strategie genannt wurden.

Aus dieser Gruppe stammen auch die meisten Patienten, deren Proben auf eigenen Wunsch oder auf Veranlassung der Urologen am Schwerpunkt Cytopathologie weiter untersucht werden; man kann ihren Anteil mit ca. 80% angeben. Hinzu kommen wenige Patienten, die die *Active Surveillance*-Kriterien sicher erfüllen und sich mehr Sicherheit bei der Wahl ihrer Therapieoption versprechen, sowie wenige Patienten, die trotz prospektiver Aussichtslosigkeit einer *Active Surveillance*-Strategie (z.B. bei einem Gleason-Grad 9 und einem PSA von 24 ng/ml) einen komplementären Rat suchen. Gerade bei letzterer Gruppe werden Messungen nur nach Rücksprache mit dem behandelnden Urologen durchgeführt, weil im Sinne des Wirtschaftlichkeitsgebotes in der Medizin hier formal eine Untersuchung vorgenommen wird, deren Indikationsstellung sich dem Pathologen nicht eigenständig erschließt.

Was bei Konzeption unserer methodischen Studie allerdings noch fehlte, war eine Untersuchung zur Validierung der Empfehlungen zur *Active Surveillance*-Strategie am eigenen Krankengut. Letzten Endes erfolgten die klinischen Wertungen der zugrunde liegenden Messungen auf der Basis der Arbeit von Tribukait (1993), die in Kapitel 1.4 dargestellt und deren Hauptergebnis in Abb. 4 wiedergegeben ist. Ziel der Arbeit war aber nicht, etwas über *Active*

Surveillance auszusagen, sondern allgemein den Einfluss DNA-bildzytometrischer Kenngrößen auf den Langzeitverlauf von Patienten mit Prostatakarzinomen ohne Behandlung zu untersuchen, so dass die Arbeit sowohl *Active Surveillance*-Patienten, als auch Patienten mit dem Konzept des *Watchful Waiting* sowie ggfls. auch Patienten mit anderen Gründen für eine „Nicht-Behandlung“ zusammenfasst. Inzwischen liegen aber Validierungsdaten aus dem Schwerpunkt Cytopathologie und dem Institut für Pathologie in Düren vor, die zur Veröffentlichung eingereicht sind (Böcking et al. 2017) und die die DNA-bildzytometrischen prognostischen Studien zum Prostatakarzinom, die in einem systematischen Review zusammengefasst wurden, ergänzen werden (Böcking et al. 2014).

Bezüglich der interspezifischen Reproduzierbarkeit der DNA-Bildzytometrie sind auf der Basis von Untersuchungen diverser Arbeitsgruppen die oben bereits erwähnten Konsensus-Publikationen entwickelt worden (Böcking et al. 1995, Giroud et al. 1998, Haroske et al. 1998, Haroske et al. 2001), die eine sehr gute Verlässlichkeit der DNA-bildzytometrischen Messtechnik belegen. Ferner zeigen die aktuellen Auswertungen von Tils, dass die Zuordnung von DNA-bildzytometrischen Mustern zu der Klassifikation von Tribukait mit hoher Reproduzierbarkeit möglich ist (Tils 2014). Aus alledem resultierte auch die 2015 erschienene Veröffentlichung „Objektive DNA-Malignitätsgradierung als Ergänzung zum histologischen Gleason-Score“, in der alle deutschsprachigen Arbeitsgruppen, die mit der DNA-bildzytometrischen Diagnostik des Prostatakarzinoms befasst sind, aktuell ihre gemeinsame Vorgehensweise dokumentiert haben (Böcking et al. 2015).

Wenn nun die Voraussetzungen für eine technisch einwandfreie Anwendung der DNA-Bildzytometrie als erfüllt gelten können, so stellt sich die Frage, weswegen die Methode auf wenige diagnostische Schwerpunkte beschränkt geblieben ist. Eine der Antworten hierauf ist die relativ lange Bindung von Personal zur klassischerweise manuellen Messung, die bislang die Aufmerksamkeit eines geschulten Operators während des gesamten Messvorganges und zu seiner Auswertung voraussetzt. Entsprechend wurden Möglichkeiten

gesucht, die eigentliche Messung vollautomatisch auszuführen und die Tätigkeit des Operators auf Kontrollaufgaben und Supervision zu beschränken.

5.4 Automatisierte DNA-Bildzytometrie - Möglichkeiten und Grenzen

Die Erprobung einer solchen Automatisierung war Gegenstand der hier dargestellten Studie. Dabei beschränkte sich diese auf die reine Anwendung einer ersten lauffähigen Softwareversion, die aus einer längerfristigen Beschäftigung mit Problemen der Zellerkennung und Zellzuordnung resultierte.

Wie in der Einleitung bereits dargestellt, wurden dazu zunächst Tausende von Bildern von Zellkernen aus Zellvereinzelungspräparaten von Prostatakarzinomen durch einen erfahrenen Zytopathologen den verschiedenen Zellarten zugeordnet. Aufgabe der Kooperationspartner an der RWTH Aachen war dann, geeignete Eigenschaften pro Zellart zu extrahieren, um zu einer möglichst guten automatisierten Zuordnung der Zellen zu gelangen und so die Generierung von Klassifikatoren zu ermöglichen. Grundprinzipien der Vorgehensweise sind in der bereits erwähnten Promotionsschrift von Friedrich (2015) dargestellt, eine grundlegende Tabelle mit Klassifikationsergebnissen wurde in unserer Arbeit in Abb. 6 wiedergegeben. Man erkennt, dass rein auf der Basis von fast genau 48.000 Zellbildern eine Klassifikationsgüte von knapp 91% erreicht werden konnte. Fehldetektionen betrafen dabei fast jede mögliche Kombination von Zellen bzw. Zellinterpretationen.

Für den wirklichen Messvorgang kam es allerdings auf andere Dinge bzw. mit auf andere Dinge an. Eine komplett richtige Zuordnung der Zellen zu den einzelnen Zellarten ist nämlich im Grunde gar nicht verlangt, solange man die Gewähr dafür übernehmen kann, dass man die eigentlichen Messparameter verlässlich erheben kann.

Beispielsweise ist egal, wie viele Zellen sich innerhalb einer bestimmten DNA-Stammlinie befinden, sofern die Lage der DNA-Stammlinie stimmt. Gleiches gilt auch für Variable wie den 9cEE: Kommt es nur darauf, ob solche Zellen

vorhanden sind (also $9cEE = 0$ vs. $9cEE > 0$), so wäre theoretisch ausreichend, wenn eines der Objekte, die oberhalb von $9cEE$ eingeordnet würden, wirklich einer Analysezellen entspräche; ob die anderen Cluster oder anderweitige Artefakte wären, könnte letztendlich als irrelevant angesehen werden. Auch fehlgemessene, möglicherweise defokussierte Eichzellen wären nicht störend, sofern sich der Peak der Eichzellen so bestimmen ließe, dass eine normale Zellpopulation eine DNA-Stammlinie bei $2c$ aufweisen würde.

In der Realität einer wissenschaftlichen Studie geht man selbstverständlich anders vor und versucht, ein optimales Ergebnis für alle Zellarten zu erreichen. Das aber wiederum kostet eine gewisse Zeit, und man kann sagen, dass bezogen auf die hier zur Verfügung stehende Softwareversion keine kürzere Bearbeitungszeit für den Untersucher anfiel als bei einer entsprechenden manuellen Messung. Dies bedeutet, dass das Ziel einer Zeitersparnis mit dieser Softwareversion, die aber heute bereits überholt ist, nicht erreichbar war.

Eine große Zahl von Objekten, die aber nicht erneut quantifiziert wurden, war aus der eigenen Messerfahrung heraus richtig zugeordnet. Fehlzugeordnungen durch den *Classifier* betrafen vor allem folgende beiden Bereiche:

- Zuordnung von sich überschneidenden Analysezellen als eine einzelne Analysezeile statt als Artefakt
- Zuordnung von eindeutig defokussierten Elementen als Referenz- oder Analysezeile statt als Artefakt

Die sich überschneidenden Analysezellen kann man sich in vielen Fällen als zweidimensionale Projektion einer Weinrebe vorstellen, also als ein Konglomerat von Objekten, wobei jede einzelne Traube eine Analysezeile repräsentiert (vgl. Abb. 16). Die Erwartung, man könne Parameter entwickeln, um solche rebenförmigen Konglomerate anhand ihrer Größe oder ihrer „unrunden“ Form mit dem Auf und Ab von Kreis- oder Ellipsenbögen zu eliminieren, hat sich bislang nicht optimal umsetzen lassen - jedenfalls fanden sich solche Konglomerate nicht nur in den Artefakten (richtige Zuordnung), sondern auch in der Klasse der Analysezellen (falsche Zuordnung). Das

Problem ist dabei bereits aus mathematisch-bildanalytischer Sicht nicht ganz trivial, wie Diskussionen mit den beteiligten Naturwissenschaftlern der RWTH Aachen zeigten. Auch aus medizinischer Sicht stellt es sich komplex dar: Zum einen muss gewährleistet sein, dass man nicht durch das Setzen einer Obergrenze für die Zellkernfläche, oberhalb derer automatisch eliminiert würde, auf einmal wirkliche Analysezellen mit besonders großen Kernen ausgrenzte - dies könnte das Messergebnis inhaltlich im Einzelfall verfälschen. Außerdem muss auch der Tatsache Rechnung getragen werden, dass Tumorzellkerne durchaus bizarre Formen annehmen können, also durchaus deutlich von einer typischen geometrischen Grundfigur (bzw. auf einem Bild von einer zweidimensionalen Projektion eines solchen Körpers) abweichen können. Dennoch fiel aber auf, dass ein großer Teil der Eliminationen von Objekten aus der Gruppe der Analysezellen auf solchen sich überschneidenden Objektkonglomeraten (Abb. 16) oder Dubletten (Abb. 17) etc. bestand, von denen man aus morphologischer Sicht angenommen hätte, dass ihre automatische Zuordnung als Artefakt möglich gewesen wäre. Hier wäre also ggfls. noch Spielraum für Softwareverbesserungen.

Zu überlegen wäre auch, ob es nicht Sinn machen könnte, Objekte mit „zu hohen“ DNA-Werten generell zu löschen. Die Erfahrung am Schwerpunkt Cytopathologie aus mehr als eintausend manuellen Messungen lehrt, dass Zellen oberhalb von $32c$, also oberhalb von 2^5c , d.h. jenseits der G_2/M''' -Phase, als extrem selten gelten können, abgesehen vielleicht von Geweben mit euploider Polyploidisierung (Biesterfeld et al. 1994): Bei diesen sind erheblich höhere Ploidiestufen möglich, z.B. beim Herzmuskel und beim Samenblasenepithel, und auch der höchste jemals in Düsseldorf gemessene Einzelwert von $128c$ wurde in einem polyploiden Gewebe, nämlich der Schilddrüse, nachgewiesen (Heide 1983). Mit einer solchen Streichung ab $32c$ würde sich im Einzelfall manche Elimination von Artefakten von selbst erledigen, ohne dass ein Risiko für eine Ergebnisverfälschung resultierte.

Defokussierte Elemente werden üblicherweise mit zu geringen IOD-Werten gemessen, so dass der angegebene DNA-Einzelwert dann zu niedrig liegt. In unserer Softwareversion war keine automatische Zuordnung von Objekten zu den Artefakten vorgesehen, die einen bestimmten DNA-Mindestwert

unterschreiten. Stattdessen war die Software rein darauf ausgelegt, defokussierte Elemente anhand morphometrischer Parameter zu erkennen. Dieses gelang aber nicht immer in ausreichendem Maße, so dass eine Reihe defokussierter Objekte nachträglich umklassifiziert werden musste. Ein Beispiel für ein solches Objekt ist vorne in Abb. 18 gezeigt. Indikatoren wie eine ausgefranste Kernkontur oder zentral abgeblasstes Chromatin sowie das für das menschliche Auge eindeutige „Verstreichen“ der Kernkontur sind sicher gute Indikatoren für Defokussierung. Es gilt aber auch hier, dass verhindert werden muss, dass Messobjekte im Fokus, die ähnliche Eigenschaften aufweisen, fälschlich von vornherein aus der Messung ausgeschlossen werden. Die Empfehlung, unterhalb eines bestimmten DNA-Mindestwertes aller Zellen zu verwerfen, ist im Prinzip insofern sinnvoller, als nur sehr selten eine Messung auftreten wird, in der viele Tumorzellen hypodiploid sein werden, also DNA-Gehalte von signifikant weniger als $2c$ aufweisen. Ein Beispiel hierfür wären die seltenen chromophoben Nierenzellkarzinome, bei denen chromosomale Verluste üblich sind. Als sinnvoller DNA-Mindestwert käme eine Schwelle von $1,75c$ in Betracht.

Ein weiterer wesentlicher Kritikpunkt betrifft den Autofokus des Systems, und zwar nicht bezogen auf die Messung eines einzelnen defokussierten Objektes, sondern im Ganzen. Durch das Setzen von sogenannten Fokuspunkten (siehe Kap.3.5) gibt der Untersucher der Software vor, in welchem Bereich der Z-Achse der Autofokus für den Fall, dass in der Umgebung des Fokuspunktes Objekte erkannt werden sollen, arbeiten soll. Hier zeigte das Meßsystem in unserer Softwareversion noch Schwächen, da sich gelegentlich eine nicht mehr rückführbare Defokussierung ergab, hervorgerufen dadurch, dass die Fokusebene unkorrigierbar außerhalb der Zone lag, in welcher der Autofokus auf der Z-Achse arbeitete. Ggfls. wäre hier auch etwas gewonnen, wenn die Mikroskopeinrichtung besser druckgepuffert wäre und auf diese Weise weniger äußeren Einflüssen wie Erschütterungen ausgesetzt wäre.

Klar ist, dass die messtechnischen Schwierigkeiten auch bei der Erhebung der Messgrößen wie Lage der DNA-Stammlinien etc. fortwirken. Sie haben also nicht nur eine formale, sondern auch eine klinische Bedeutung, die es zu

bedenken gibt. Es wäre somit auch aus diagnostischer Sicht wünschenswert, wenn Weiterentwicklungen die beschriebenen Schwächen des Meßsystems weiter verringern würden (s. unten).

Dieses ist aber nicht nur unter messtechnischen, sondern auch unter klinischen Gesichtspunkten wünschenswert. Auch wenn sich (s. Kapitel 4.4.2) zwischen der manuellen und der automatisierten Messung eine Übereinstimmung im Ploidie-Typ nach Tribukait von 81,5% (22 von 27 Fällen) ergab und der Kappa-Wert von 0,622 als „gute“ Übereinstimmung gilt, stellt sich dennoch das Problem, dass fünf Patienten nicht ausreichend sinnvoll hätten beraten werden können, wenn beide Messergebnisse parallel vorgelegen hätten. Auf der Basis der manuellen Messung jedenfalls konnte man zwei der fünf Patienten (Tribukait C) keine *Active Surveillance*-Strategie empfehlen, während bei dreien zumindest bedingt ein solcher Weg zu diskutieren gewesen wäre (Tribukait B). Leider lagen uns weder Daten zu der abschließend getroffenen klinischen Therapieentscheidung noch klinische Verlaufsdaten zu den Patienten vor. Es muss aber klar sein, dass weitere Verbesserungen an der Software bei klinischen Vergleichsmessungen zu einer höheren Korrelation mit der tradierten manuellen Messung, die gewissermaßen als Goldstandard dient, führen müssen.

In der Tat hat sich seit der Durchführung unserer Untersuchungen, die aus dem Jahr 2015 stammen, bereits vieles verändert. Manche Erfahrungen, die aus dieser Studie und aus anderen Projekten rückgemeldet wurden, konnten in neuen Softwareversionen berücksichtigt werden. Außerdem wurden die Geräteeinstellungen weiter optimiert. Heute ist das identische Meßsystem auch mit verbesserter Software im diagnostischen Einsatz am Institut für Pathologie in Düren und am hiesigen Schwerpunkt Cytopathologie, insbesondere an sogenannten *Monolayern*, also an speziellen Ausstrichpräparationen, die im eigenen Labor getätigt werden. Auch die Zellvereinzlungspräparate von Prostatakarzinomen gehören dazu, außerdem Pleura-, Peritoneal- und Perikardergüsse. Bei Präparaten, die bereits als von den klinischen Kollegen primär durch Ausstreichen von gewonnenem Zellmaterial hergestellt werden, also vor allem Abstriche aus der Mundhöhle, vom Gebärmutterhals oder aus

den Pankreas- und Gallengängen, ist die manuelle DNA-Bildzytometrie bislang das eher eingesetzte Verfahren, da dort oft komplexe Überlagerungen von Zellen präparatorisch nicht vermieden werden können. Entsprechend haben beide Methoden ihre besonderen Felder in der adjuvanten zytopathologischen Diagnostik. Ob die automatisierte DNA-Bildzytometrie auch für die primäre zytopathologische Diagnostik, z.B. von Präparaten von Körperhöhlenergüssen, geeignet wäre, müsste in eigenen Versuchsreihen entschieden werden. Für die in unserer Studie untersuchten Prostatakarzinome käme eine solche Anwendung naturgemäß allerdings nicht in Betracht, da die eigentliche Primärdiagnostik dieser Tumorentität rein histologisch, nicht aber zytologisch erfolgt.

6 Literaturverzeichnis

1. Adolfsson J (2008): Watchful waiting and active surveillance: the current position. *BJU Int* 102 (1):10-14
2. Al Otaibi M, Ross P, Fahmy N, Jeyaganth S, Trottier H, Sircar K, Bégin LR, Souhami L, Kassouf W, Aprikian A, Tanguay S (2008): Role of repeated biopsy of the prostate in predicting disease progression in patients with prostate cancer on active surveillance. *Cancer* 113 (2): 286-292
3. Auer GU, Caspersson TO, Wallgren AS (1980): DNA content and survival in mammary carcinoma. *Anal Quant Cytol* 2(3): 161-165
4. Auffermann W, Böcking A (1985): Early detection of precancerous lesions in dysplasias of the lung by rapid DNA image cytometry. *Zytologie* 7 (3): 218-226
5. Bell KJL, del Mar C, Wright G, Dickinson J, Glasziou P (2015): Prevalence of incidental prostate cancer: A systematic review of autopsy studies. *Int J Cancer* 137: 1749-1757
6. Beske F, Becker E, Katalinic A, Krauss C, Pritzkeleit R (2007): Gesundheitsversorgung 2050 - Prognose für Deutschland und Schleswig-Holstein. Schriftenreihe des Fritz Beske-Institutes für Gesundheits-System-Forschung, Kiel, Band 108, ISBN 978-3-88312-335-6
7. Biesterfeld S (2014): Zurverfügungstellung urheberrechtlich geschützter Abbildungen aus dem Bestand des Funktionsbereiches Cytopathologie der Heinrich Heine-Universität Düsseldorf
8. Biesterfeld S (2015): Dokumentationsqualität histopathologischer Befundberichte von Prostatastanzbiopsien - eine Momentaufnahme. *Urologe* 53: 1644-1650
9. Biesterfeld S, Leitloff M, Rath W, Schröder W (2001a): DNA image cytometry in the differential diagnosis of endometrial hyperplasia and endometrial adenocarcinoma. *Analyt Quant Cytol Histol* 23: 123-128
10. Biesterfeld S, Deacu L (2009): DNA image cytometry in the differential diagnosis of benign and malignant lesions of the bile duct, the pancreatic duct and the papilla of Vater. *Anticancer Res* 29 (5): 1579-1584
11. Biesterfeld S, Gerres K, Fischer-Wein G, Böcking A (1994): Polyploidy in non-neoplastic tissues. *J Clin Pathol* 47: 38-42
12. Biesterfeld S, Pomjanski N, Schramm M (2013): Adjuvante Methoden in der diagnostischen Zytopathologie. *Verh Dtsch Ges Zyt* 28: 8-16
13. Biesterfeld S, Reus K, Bayer-Pietsch E, Mihalcea AM, Böcking A (2001b): DNA image cytometry in the differential diagnosis of endocervical adenocarcinoma. *Cancer* 93: 160-164

14. Böcking A, Biesterfeld S, Dietz J, Haroske G, Kriegsmann J, Motherby H, Falk S (2015): Objektive DNA-Malignitätsgradierung als Ergänzung zum histologischen Gleason-Score. *Frankfurter Konsens. Pathologe* 36 (5): 498-502
15. Böcking A, Dietz J (2013): Prognostische DNA-Zytometrie beim Prostatakarzinom. *Deutsche Zeitschrift für Onkologie* 4 (45): 2-8
16. Böcking A, Friedrich D, Börgermann C, Biesterfeld S, Engers R, Dietz J (2017): Prediction of non-progression in prostate cancer patients under active surveillance by DNA-Karyometry. *SM J Urol* 3 (1): 1030
17. Böcking A, Friedrich D, Jin C, Bell A, Würflinger T, Meyer-Ebrecht D, Zhang Y, Biesterfeld S, Pomjanski N, Bernyskyy L, Yeung R, Handschel J, Palcic B, Aach T (2013): Diagnostic Cytometry. In: Mehrotra R (Hrsg.): *Oral cytology. Kapitel 9, S. 125-144, Springer Verlag*
18. Böcking A, Giroud F, Reith A (1995): Consensus report of the European Society for Analytical Cellular Pathology task force on standardization of diagnostic DNA image cytometry. *Anal Quant Cytol Histol* 17: 1-7
19. Böcking A, Tils M, Schramm M, Dietz J, Biesterfeld S (2014): DNA-cytometric grading of prostate cancer. Systematic review with descriptive data analysis. *Pathology Discovery* 2: 7
<http://www.hoajonline.com/journals/pdf/2052-7896-2-7.pdf>
20. Bruner DW, Moore D, Parlanti A, Dorgan J, Engström P (2003): Relative risk of prostate cancer for men with affected relatives: systematic review and meta-analysis. *Int J Cancer* 107 (5):797-803
21. Burchardt M, Engers R, Müller M, Burchardt T, Willers R, Epstein JI, Ackermann R, Gabbert HE, de la Taille A, Rubin MA (2008): Interobserver reproducibility of Gleason grading: evaluation using prostate cancer tissue microarrays. *J Cancer Res Clin Oncol.* 134(10): 1071-1078
22. Cheng HH, Pritchard CC, Montgomery B, Lin DW, Nelson PS (2017): Prostate Cancer Screening in a New Era of Genetics. *Clin Genitourin Cancer*, Epub ahead of print
23. Cooner WH, Mosley BR, Rutherford CL, Beard JH, Pond HS, Terry WJ, Igel TC, Kidd DD (1990): Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate specific antigen. *J Urol* 143 (6): 1146-1152
24. Damjanoski I, Müller J, Schnöller TJ, Kufer R, Rinnab L (2013): Histo-pathologische Befundberichte von Prostatastanzbiopsien. Individuelle Behandlungskonzepte. *Urologe* 52 (2): 226-239
25. Delgado R, Mikuz G, Hofstädter F (1984): DNA-Feulgen-cytophotometric analysis of single cells from paraffin embedded tissue. *Path Res Pract* 179: 92-94

26. Desai A, Wu H, Sun L, Sesterhenn IA, Mostofi FK, McLeod D, Amling C, Kusuda L, Lance R, Herring J, Foley J, Baldwin D, Bishoff JT, Soderdahl D, Moul JW (2002): Complete embedding and close step-sectioning of radical prostatectomy specimens both increase detection of extraprostatic extension, and correlate with increased disease-free survival by stage of prostate cancer patients. *Prostate Cancer Prostatic Dis* 5 (3):212-218
27. Dutch Urological Association: Prostate Cancer. Nation-wide guideline. Version 1.0, Maastricht, 2007
28. Eaton NE, Reeves GK, Appleby PN, Key TJ (1999): Endogenous sex hormones and prostate cancer: a quantitative review of prospective studies. *Br J Cancer* 80(7): 930-934
29. Egevad L, Mazzucchelli R, Montironi R (2012): Implications of the International Society of Urological Pathology modified Gleason grading system. *Arch Pathol Lab Med* 136: 426-434
30. Epstein JI, Allsbrook WC, Amin MB, Egevad LL (2005): The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. *Am J Surg Pathol* 29: 1228-1242
31. Feulgen R., Rossenbeck H (1924): Mikroskopisch-chemischer Nachweis einer Nukleinsäure vom Typus der Thymonukleinsäure und die darauf beruhende Elektiv Färbung von Zellkernen in mikroskopischen Präparaten. *Z Physiol Chem* 135: 203-248
32. Fleischer M (2017): DNA-Bildzytometrie als prognostischer Marker bei Borderlinetumoren des Ovars. Med. Diss., Heinrich Heine-Universität Düsseldorf
33. Fornara P, Fischer K, Luboldt H, Döhn C, Semjonow A (2004): Einfluss von Präanalytik und Analytik auf die Aussagekraft des prostata-spezifischen Antigens. *Deutsches Ärzteblatt* 101 (25): A 1820-1823
34. Friedrich D (2015): Effective improvement of cancer diagnostics and prognostics by computer-assisted cell image analysis. Math.-Nat. Diss., RWTH Aachen
35. Friedrich D, Chen J, Zhang Y, Demin C, Yuan L, Berynskyy L, Biesterfeld S, Aach T, Böcking A (2012). Identification of Prostate Cancer Cell Nuclei for DNA-Grading of Malignancy. *Bildverarbeitung für die Medizin*, Springer Verlag, Berlin Heidelberg: 334-339
36. Girmes G (2009): Stellenwert der DNA-Bildzytometrie in der Differentialdiagnostik leiomyogener Tumoren. Med. Diss., Johannes Gutenberg-Universität Mainz
37. Giroud F, Haroske G, Reith A, Böcking A (1998): 1997 ESACP consensus report on diagnostic DNA image cytometry. Part II: Specific recommendations for quality assurance. *European Society for Analytical Cellular Pathology. Anal Cell Pathol* 17(4): 201-208

38. Graumann W (1953): Zur Standardisierung des Schiff'schen Reagens. *Z. Wiss Mikrosc* 61: 225-226
39. Haroske G, Baak JP, Danielsen H, Giroud F, Gschwendtner A, Oberholzer M, Reith A, Spieler P, Böcking A (2001): Fourth updated ESACP consensus report on diagnostic DNA image cytometry. *Anal Cell Pathol* 23 (2): 89-95
40. Haroske G, Giroud F, Reith A, Böcking A (1998): 1997 ESACP consensus report on diagnostic DNA image cytometry. Part I: Basic considerations and recommendations for preparation, measurement and interpretation. European Society for Analytical Cellular Pathology. *Anal Cell Pathol* 17 (4): 189-200
41. Hedley DW, Friedlander ML, Taylor IW (1985): Application of DNA flow cytometry to paraffin-embedded archival material for the study of aneuploidy and its clinical significance. *Cytometry* 6: 327-333
42. Heide W (1993): Die Polyploidie der Zellkerne in Leber, Schilddrüse und Samenblasenepithel des Menschen während der 9. und 10. Lebensdekade und sie beeinflussende Faktoren. *Med. Diss., Universität Düsseldorf*
43. Heidenreich A, Aus G, Bolla M, Joniau S, Matveev V, Schmid HP, Zattoni F, European Association of Urology (2008): EAU guidelines on prostate cancer. *Eur Urol* 53 (1): 68-80
44. Helpap B, Egevad L (2006): The significance of modified Gleason grading of prostatic carcinoma in biopsy and radical prostatectomy specimens. *Virchows Arch* 449: 622-627
45. Helpap B, Egevad L (2007): Die Wertigkeit des 2005 modifizierten Gleason-Gradings in der urologischen Diagnostik von Prostatakarzinomen. *Urologe* 46: 59-62
46. Hemminki K, Czene K (2002): Age specific and attributable risks of familial prostate carcinoma from the family-cancer database. *Cancer* 95 (6):1346-1353
47. Hoogendam A, Buntinx F, de Vet HC (1999): The diagnostic value of digital rectal examination in primary care screening for prostate cancer: a meta-analysis. *Fam Pract* 16(6): 621-626
48. Khatami A, Aus G, Damber JE, Lilja H, Lodding P, Hugosson J. (2007): PSA doubling time predicts the outcome after active surveillance in screening-detected prostate cancer: results from the European randomized study of screening for prostate cancer, Sweden section. *Int J Cancer* 120(1): 170-4 <http://www.ncbi.nlm.nih.gov/pubmed/17013897>.
49. Klotz L (2005): Active surveillance for prostate cancer: for whom? *J Clin Oncol* 23 (32): 8165-8169

50. Klotz L, Zhang L, Lam A, Nam R, Mamedov A, Loblaw A (2010): Clinical results of long-term follow up of a large, active surveillance cohort with localized prostate cancer. *J Clin Oncol* 28(1): 126-131,
51. Krishnamurthy S, Katz RL, Shumate A, Strohle K, Khanna A, Tucker SL, Rajman I, Lahoti S (2001): DNA image analysis combined with routine cytology improves diagnostic sensitivity of common bile duct brushing. *Cancer* 93: 229-235
52. Kuroiwa K, Shiraishi T, Ogawa O, Usami M, Hirao Y, Naito S for the Clinicopathological Research Group for Localized Prostate Cancer Investigators (2010): Discrepancy between local and central pathological review of radical prostatectomy specimens. *J Urol* 183: 952-957
53. Lavoipierre AM, Snow RM, Frydenberg M, Gunter D, Reisner G, Royce PL, Lavoipierre GJ (1998): Prostatic cancer: role of color Doppler imaging in transrectal sonography. *Am J Roentgenol* 171 (1): 205-210
54. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF): Konsultationsfassung: Interdisziplinäre Leitlinie der Qualität S3 zur Früherkennung, Diagnose und Therapie der verschiedenen Stadien des Prostatakarzinoms, Langversion 4.0, 2016. AWMF Registernummer: 043/022OL
<http://leitlinienprogramm.onkologie.de/Prostatakarzinom.58.0.html> (Zugriff am: 09.05.2017)
55. Melia J, Moseley R, Ball RY, Griffiths DD, Grigor K, Harnden P, Jarmulowicz M, McWilliam LJ, Montironi R, Waller M, Moss S, Parkinson MC (2006): A UK-based investigation of inter- and intraobserver reproducibility of Gleason grading of prostatic biopsies. *Histopathology* 48 (6): 644-654
56. Motherby H, Marcy T, Hecker M, Ross B, Nadjari B, Auer H, Müller KM, Häussinger D, Strauer BE, Böcking A (1998): Static DNA cytometry as a diagnostic aid in effusion cytology: I. DNA aneuploidy for identification and differentiation of primary and secondary tumors of the serous membranes. *Anal Quant Cytol Histol* 20: 153-161
57. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, Fossati N, Gross T, Henry AM, Joniau S, Lam TB, Mason MD, Matveev VB, Moldovan PC, van den Bergh RC, van den Broeck T, van der Poel HG, van der Kwast TH, Rouvière O, Schoots IG, Wiegel T, Cornford P (2017): EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* 71 (4):618-629
58. Netto G J, Eisenberger M, Epstein JI for the TAX 3501 Trial Investigators (2011): Interobserver variability in histologic evaluation of radical prostatectomy between central and local pathologists: Findings of TAX 3501 Multinational Clinical Trial. *Urology* 77: 1155-1160

59. NICE (National Institute for Health and Clinical Excellence). Prostate Cancer: diagnosis and treatment. Full guideline, 2008
https://www.ncbi.nlm.nih.gov/books/NBK49533/pdf/Bookshelf_NBK49533.pdf
60. Osterheld MC, Andrejevic Blant S, Caron L, Braunschweig R, Dorta G, Bouzourene H, Mihaescu A (2005): Digital image DNA cytometry: a useful tool for the evaluation of malignancy in biliary strictures. *Cell Oncol* 27 (4): 255-260
61. Paul R, Zimmermann F, Dettmar P, Adam M, van Randeborgh H, Alschibaka M, Herkommer K, Seitz M: Prostatakarzinom. In: Treiber U, Zaak D (Hrsg.): Manual "Urogenitale Tumoren", Tumorzentrum München und W. Zuckschwerdt Verlag GmbH, Germering, 2008
62. Robert Koch-Institut (Hrsg.) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg.): Krebs in Deutschland 2005/2006. Häufigkeiten und Trends. 7. Ausgabe, Berlin, 2010
63. Robert Koch-Institut (Hrsg.) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg.): Krebs in Deutschland 2007/2008. 8. Ausgabe, Berlin, 2012
64. Robert Koch-Institut (Hrsg.) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg.): Krebs in Deutschland 2011/2012. 10. Ausgabe, Berlin, 2015
65. Rohde V, Katalinic A, Wasem J (2007): Prostataerkrankungen. In: Robert-Koch Institut (Hrsg.): Gesundheitsberichterstattung des Bundes, Heft 36
66. Rubin MA, Bismar TA, Curtis S, Montie JE (2004): Prostate needle biopsy reporting: How are the surgical members of the Society of Urologic Oncology using pathology reports to guide treatment of prostate cancer patients?. *Am J Surg Pathol* 28 (7): 946-952
67. Sakr WA, Tefilli MV, Grignon DJ, Banerjee M, Dey J, Gheiler EL, Tiguert R, Powell IJ, Wood DP (2000): Gleason score 7 prostate cancer: a heterogeneous entity? Correlation with pathologic parameters and disease-free survival. *Urology* 56 (5): 730-734
68. Sandritter W, Carl M, Ritter W (1966): Cytophotometric measurements of the DNA content of human malignant tumors by means of the Feulgen reaction. *Acta Cytol* 10(1): 26-30
69. Seidel A, Sandritter W (1963): Cytophotometric Measurements of the DNA Content of a Pulmonary Adenoma and of a Malignant Pulmonary Adenomatosis. *Z Krebsforsch* 65: 555-559
70. Shappley WV, Kenfield SA, Kasperzyk JL, Qiu W, Stampfer MJ, Sanda MG, Chan JM (2009): Prospective study of determinants and outcomes of deferred treatment or watchful waiting among men with prostate cancer in a nationwide cohort. *J Clin Oncol* 27 (30): 4980-4985

71. Tils M (2016): Verteilung von DNA-Ploidiemustern in Stanzbiopsien von Prostatakarzinomen. Med. Diss., Heinrich Heine-Universität Düsseldorf.
72. Tribukait B (1993): Nuclear deoxyribonucleic acid determination in patients with prostate carcinomas: clinical research and application. *Eur Urol* 23 Suppl 2: 64-76
73. Uemura H, Hoshino K, Sasaki T, Miyoshi Y, Ishiguro H, Inayama Y, Kubota Y (2009): Usefulness of the 2005 International Society of Urologic Pathology Gleason grading system in prostate biopsy and radical prostatectomy specimens. *BJU Int* 103 (9): 1190-1194
74. Vainer B, Toft B G, Olsen K E, Jacobsen G K, Marcussen N (2011): Handling of radical prostatectomy specimens: total or partial embedding?. *Histopathology* 58 (2): 211-216
75. Veloso SG, Lima MF, Salles PG, Berenstein CK, Scalon JD, Bambirra EA (2007): Interobserver agreement of Gleason score and modified Gleason score in needle biopsy and in surgical specimen of prostate cancer. *Int Braz J Urol* 33 (5): 639-646
76. Wilt T J, MacDonald R, Rutks I, Shamliyan T A, Taylor B C, Kane R L (2008): Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer. *Ann Intern Med* 148 (6): 435-448
77. Zareba P, Zhang J, Yilmaz A, Trpkov K (2009): The impact of the 2005 International Society of Urological Pathology (ISUP) consensus on Gleason grading in contemporary practice. *Histopathology* 55 (4): 384-391

Danksagung

Ich möchte mich bei allen bedanken, die mich bei der Promotion begleitet und unterstützt haben.

Ein ganz besonderer Dank gilt meinem Doktorvater Prof. Dr. med. Stefan Biesterfeld. Nur durch seine aktive und engagierte Unterstützung in allen wissenschaftlichen und organisatorischen Fragen, konnte diese Dissertation vollendet werden.

Mein Dank gilt auch Dr. med. Martin Schramm und dem ganzen Team des Instituts für Pathologie, Funktionsbereich Cytopathologie, für ihre tatkräftige Unterstützung.

Familie Dr. Haiaty möchte ich danken für ihr Interesse am Erfolg dieser Dissertation und ihre Unterstützung in allen wichtigen Lebenssituationen.

Der größte Dank gilt meiner Familie, Esmail, Farideh, Zahra und Hoda Rezaie sowie meiner Frau Hedye Haiaty, für ihre moralische Unterstützung und stetige Motivation. Ohne euch hätte ich es nicht so weit geschafft.

-Vielen Dank-