

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

Direktorin: Univ.-Prof. Dr. med. I. Esposito
Funktionsbereich Cytopathologie

**Erprobung eines automatischen DNA-bildzytometrischen
Messsystems an Schilddrüsenpunktaten**

Dissertation

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

vorgelegt von

Tugba Dogan

2023

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

gez.:

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

Erstgutachter: Prof. Dr. med. Stefan Biesterfeld

Zweitgutachter: Prof. Dr. med. Nikolas H. Stoecklein

meinem Baby gewidmet

Zusammenfassung

Mit der DNA-Bildzytometrie steht am Funktionsbereich Cytopathologie der Universitätsklinik Düsseldorf eine adjuvante TV-bildanalytische Untersuchungsmethode für den diagnostischen Alltag zur Verfügung, mit der sich über die Diagnose einer numerischen DNA-Aneuploidie die neoplastische Transformation von Zellpopulationen belegen lässt. Bisher wird diese Methode angewandt, indem ein Untersucher an einem Bildschirm die zu messenden Referenzzell- und Analysezellkerne mit einer Maus anklickt und sie so einzeln der Messung hinzufügt ("manuelle DNA-Bildzytometrie"). Ziel dieser Arbeit war, einen Prototyp eines automatisierten Messsystems für die DNA-Bildzytometrie an luftgetrockneten Schilddrüsenpunktaten zu erproben ("automatische DNA-Bildzytometrie") und einen direkten Vergleich mit früheren manuellen Messungen am identischen Ausstrich zu ziehen. Der Prototyp des automatischen Messsystems stammte vom gleichen Hersteller (Fa. Motic, Xianmen, China) wie das manuelle Messsystem auch (jeweils MotiCyte DNA Workstation) und wurde in Kooperation mit Universitätspartnern an der RWTH Aachen und der Heinrich-Heine-Universität Düsseldorf entwickelt.

Für die Untersuchung wurden DNA-bildzytometrische Messungen an 40 luftgetrockneten Ausstrichpräparaten von Feinnadelpunktaten der Schilddrüse von Fällen vorgenommen, in denen aus der Primärdiagnostik eine manuelle DNA-bildzytometrische Messung und ferner ein histologisches Follow-Up vorlagen. Die automatischen Messungen wurden nachträglich unter Verwendung eines neu entwickelten Klassifikators zur automatischen Zellkernerkennung und -zuordnung im Rahmen dieser Studie vorgenommen.

Für die manuelle Messung wurden je Präparat ca. 300 Analyse- und 30 Referenzzellen ausgewählt, bei der automatischen Messungen wurden alle Zellen des gesamten Präparates erfasst. Für die anschließende Datenauswertung wurden die Zahl der Analysezellen und der Referenzzellen, die mittlere Referenzzell-IOD, der minimale und der maximale C-Wert, der 5cEE (5cER), der 9cEE (9cER) und die Lage der DNA-Stammlinie vergleichend betrachtet. Ferner konnten Daten zur Gesamttrefferlichkeit, Sensitivität, Spezifität, positivem prädiktivem Wert und negativem prädiktivem Wert der konventionellen Zytologie sowie der manuellen und automatischen DNA-Bildzytometrie ermittelt werden.

Im methodischen Vergleich erwiesen sich die Ergebnisse beider Messreihen bezüglich der Datenverteilung und statistischer Korrelationen als im Wesentlichen gut vergleichbar. Betrachtet man beide Messreihen durch Auswertung nach Lage der DNA-Stammlinie, nach der Zahl der 9cEE und in Kombination beider Größen dahingehend, ob sie als DNA-aneuploid oder als DNA-euploid zu interpretieren sind, so zeigt sich beim Vergleich der DNA-Stammlinien eine Übereinstimmung von 80%, für den 9cEE von 85% und für die kombinierte Auswertung von 72,5%. Insgesamt ergaben sich bezüglich der Interpretation der Messungen als DNA-euploid oder DNA-aneuploid in

14 Fällen (35%) Diskrepanzen bezüglich der DNA-Stammlinie oder dem 9cEE, die sich nicht in allen Fällen aufklären ließen, aus denen sich aber Anregungen für die Weiterentwicklung der Automatisierung in der DNA-Bildzytometrie ableiten lassen. In zwei Fällen (5%) blieben die Unterschiede in den Messergebnissen unauflösbar.

Vergleicht man die Ergebnisse beider Messmethoden bezogen auf die Treffsicherheit der drei diagnostischen DNA-bildzytometrischen Parameter mit den zytologischen und den histologischen Diagnosen, so zeigt sich, dass die Treffsicherheit nach Lage der DNA-Stammlinien gut vergleichbar und nach der Zahl der 9cEE nahezu identisch waren (jedoch mit diskrepanten Fällen, die näher betrachtet wurden). Bei der Treffsicherheit der kombinierten Auswertung zeigte die manuelle Messreihe etwas bessere Kennwerte.

Die hier vorgelegte Studie basiert auf Untersuchungen an einem Prototyp mit einer inzwischen weiterentwickelten Software-Version aus dem Jahr 2011/2012, die seitdem in China fortgeschrieben worden ist. Insofern ist zu erwarten, dass sich über die vergangenen Jahre Verbesserungen in der Software ergeben haben, die sich positiv in der Anwendbarkeit der automatisierten Systeme auswirken könnten. Auf dem Stand der Software, der hier zur Verfügung stand, lässt sich sagen, dass eine Automatisierung der DNA-Bildzytometrie auch an luftgetrockneten Ausstrichpräparaten technisch möglich ist, dass aber noch Defizite in der Methodik beseitigt werden müssen, um das Ziel einer wirklich automatischen Befunderstellung, bei der ein Untersucher nur noch zur Überprüfung der Bildgalerie benötigt wird, zu erreichen. Unsere Studie kann aber als eine systematische Pilotstudie des Lernens angesehen werden, aus der sich Schlüsse für weitere Planungen der Automatisierung bzw. der Interpretation ihrer Ergebnisse ziehen lassen.

Summary

DNA image cytometry is an adjuvant TV image-analytical examination method available at the Cytopathology Division of the University Hospital Düsseldorf for everyday diagnostic use. It allows neoplastic transformation of cell populations to be proven via diagnosis of a numerical DNA aneuploidy. Until now, this method has been applied by an investigator clicking on the reference cell and analysis cell nuclei to be measured on a screen using a mouse and thus adding them to the measurement one by one ("manual DNA image cytometry"). The aim of this work was to test a prototype of an automated measurement system for DNA image cytometry on air-dried thyroid punctates ("automated DNA image cytometry") and to draw a direct comparison with earlier manual measurements on the identical smear. The prototype for the automated measuring system came from the same manufacturer (Motic, Xianmen, China) as the manual measuring system (MotiCyte DNA Workstation) and was developed in cooperation with university partners at RWTH Aachen and Heinrich- Heine- University Düsseldorf.

For the study, DNA image cytometric measurements were carried out on 40 air-dried smear preparations from fine needle punctates of the thyroid gland from cases in which a manual DNA image cytometric measurement as well as a histological follow-up were available from the primary diagnostics. The automated measurements were subsequently carried out using a newly developed classifier for automated cell nucleus recognition and assignment within the scope of this study.

For the manual measurement, approximately 300 analysis cells and 30 reference cells were selected per preparation; for the automated measurements, all cells of the entire preparation were recorded. For the subsequent evaluation of the data, the number of analysis and reference cells, the mean reference cell IOD, the minimum and maximum C-value, the 5cEE (5cER), the 9cEE (9cER) and the position of the DNA stem line were considered comparatively. Furthermore, data on the overall accuracy, sensitivity, specificity, positive and negative predictive value of conventional cytology as well as manual and automated DNA image cytometry were determined.

In the methodological comparison, the results of both series of measurements proved to be essentially comparable with regard to data distribution and statistical correlations. We also looked at both series of measurements by evaluating them according to the position of the DNA stem line, according to the number of 9cEE and in combination of both variables to determine whether they are to be interpreted as DNA-aneuploid or DNA-euploid; agreement was 80% for the DNA stem lines, 85% for the 9cEE and 72.5% for the combined evaluation. In 14 cases (35%) there were discrepancies in the interpretation of the measurements as DNA euploid or DNA aneuploid with regard to the DNA stem line or the 9cEE, which could not be resolved in all cases; however, this offers suggestions for the further development of automation in DNA image cytometry. In two cases (5%), the differences in the measurement results remained unresolvable.

Comparing the results of both measurement methods related to the accuracy of the three diagnostic DNA image cytometric parameters with the cytological and histological diagnoses, it can be seen that the accuracy according to the location of the DNA stem lines was comparable and according to the number of 9cEE was almost identical (however, with discrepant cases that were investigated more closely). For the accuracy of the combined evaluation, the manual measurement series showed slightly better characteristic values.

This study is based on investigations using a prototype with a software version from 2011/2012 that has since been further developed in China. In this respect, it is to be expected that improvements have been made to the software over the past years, which could have a positive effect on the applicability of the automated systems. Based on the state of the art of the software that was available to us, it can be said that automation of DNA image cytometry is technically possible even on air-dried smear preparations, but that deficits in the methodology still need to be eliminated in order to achieve the goal of truly automated report generation, where an examiner is needed only to check the image gallery. However, our study can be seen as a systematic pilot study of learning, from which conclusions can be drawn for further planning of automation or the interpretation of its results.

Abkürzungsverzeichnis

µg	Mikrogramm
2c Dev. Index	2c Deviation Index
2c Ref. IOD	2c Referenz IOD
99mTc-Pertechnetat	99m-Technetium-Pertechnetat
Abb.	Abbildung
AK	Antikörper
aqua dest.	aqua destillata, destilliertes Wasser
ATC	Anaplastic Thyroid Carcinoma
AUC	area under curve
auto.	automatisch/automatisiert
BRAF	proto-oncogene B-Raf
c	content, Einheit für den DNA-Gehalt
CC	coefficient of correlation = Korrelationskoeffizient
CCD Kamera	Charge Coupled Device Kamera
CK-19	Cytokeratin-19
CT	Computertomographie
CTA	Chemisch-Technische Assistenz
CV	coefficient of variation = Variationskoeffizient
C-Zellen	C für Calcitonin
DGP	Deutsche Gesellschaft der Pathologie
DGZ	Deutsche Gesellschaft der Zytologie
DNA	desoxyribonucleic acid
EE	exceeding events
ER	exceeding rates
ESACP	European Society for Analytical Cellular Pathology
FISH	Floreszenz-in-situ-Hybridisierung
FNAB	Feinnadelaspirationsbiopsie
FNP	Feinnadelpunktion
FTC	Follicular Thyroid Carcinoma
g	Gramm
G2-Phase	Gap-2-Phase des Zellzyklus
Ggf.	gegebenenfalls
HBME	Hector Battifora mesothelial epitope-1
HCL	Salzsäure, Chlorwasserstoffsäure
HPV	Humanes Papillom Virus
Hrsg.	Herausgeber
IOD	integrierte optische Dichte
KH ₂ PO ₄	Kaliumdihydrogenphosphat
KIGGS	Studie zur Gesundheit vom Kindern und Jugendlichen in Deutschland
MEN	multiple endocrine Neoplasie
mg	Milligramm
MGG	May-Grünwald-Giemsa
MIFTC	Minimally Invasive Follicular Thyroid Cancer
ml	Milliliter
M-Phase	Mitose-Phase
MRT	Magnetresonanztomographie
MTA	Medizinisch-technische Assistenz

MTC	Medullary Thyroid Carcinoma
Na ₂ HPO ₄	Dinatriumhydrogenphosphat
NIFTP	Noninvasive Follicular Thyroid Neoplasm With Papillary-like nuclear features
NPV	negative predictive value
PAP-Test	Papanicoklaou-Test
PCR	Polymerase chain reaction
PDTC	Poorly Differentiated Papillary Thyroid Carcinoma
PPV	positive predictive value
PTC	Papillary Thyroid Carcinoma
PTEN	phosphatase and tensin homolog
RAS	rat sarcoma
RET	Rezeptor-Tyrosinkinase
RKI	Robert-Koch- Institut
RNA	Ribonukleinsäure
RWTH Aachen	Rheinisch-Westfälische Technische Hochschule
SD	Standardabweichung
SHIP	Study of Health in Pomerania
SO ₂	Schwefeldioxid
SOP	Standard Operating Procedures
S-Phase	Synthese-Phase
T3	Triiodthyronin
T4	Tetraiodthyronin
Tab.	Tabelle
TIRADS	thyreoid imaging reporting and database system
TPO	Thyreoperoxidase
TSH	Thyreoid-stimulierendes Hormon
TTF-1	Thyroidaler Transkriptionsfaktor 1
UV-Licht	Ultraviolettes Licht
WHO	World Health Organization

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Epidemiologie knotiger Veränderungen der Schilddrüse.....	1
1.2	Klassifikation von Raumforderungen der Schilddrüse.....	3
1.2.1	Benigne Schilddrüsenerkrankungen.....	3
1.2.2	Maligne Schilddrüsenerkrankungen.....	4
1.3	Präoperative Diagnostik von Schilddrüsenknoten	7
1.3.1	Anamnese und klinischer Untersuchungsbefund	7
1.3.2	Sonographie.....	7
1.3.3	Laborparameter	8
1.3.4	Bildgebende Verfahren.....	8
1.3.5	Zytologische Untersuchung.....	9
1.4	DNA-Bildzytometrie	13
1.5	Ziel der Arbeit.....	16
2	Material und Methoden.....	18
2.1	Patienten und Präparate	18
2.2	Verarbeitung der Präparate für die Routinediagnostik	18
2.3	Zytologische Diagnostik	19
2.4	Feulgen-Färbung.....	19
2.5	Geräte und Messtechnik.....	21
2.6	Prinzip der DNA-Bildzytometrie.....	22
2.7	Interaktive (manuelle) DNA-Bildzytometrie.....	22
2.7.1	Messvorgang	23
2.7.2	Datenauswertung und Dokumentation	24
2.8	Automatische DNA-Bildzytometrie	28
2.8.1	Scannen der Präparate	29
2.8.2	Digitale Klassifizierung.....	30
2.8.3	Datenauswertung und Dokumentation	31
2.8.4	Vergleichsparameter.....	37
2.8.5	Diagnostische Interpretation.....	37
2.9	Statistische Untersuchungen.....	38
2.9.1	Vier-Felder-Tafeln.....	38
2.9.2	Cohens Kappa	40
2.9.3	Bland-Altman-Diagramme	40
2.10	Ethikvotum.....	38

3	Ergebnisse.....	42
3.1	Patientenkollektiv	42
3.1.1	Ausgeschlossene Präparate.....	42
3.1.2	Alters- und Geschlechterverteilung.....	43
3.2	Morphologische Diagnostik	43
3.2.1	Zytologie	43
3.2.2	Histologie	45
3.2.3	Treffsicherheit der konventionellen Zytologie.....	45
3.3	Vergleichende Statistik zur manuellen und automatischen DNA-Bildzytometrie	46
3.3.1	Zahl der Analyse- und Referenzzellen	46
3.3.2	Mittlere Referenzzell-IOD	47
3.3.3	Minimaler C-Wert	47
3.3.4	Maximaler C-Wert	49
3.3.5	5c EE / 5cER	51
3.3.6	9c EE / 9cER	55
3.3.7	DNA-Stammlinie	60
3.4	Diagnostische Interpretation der manuellen und der automatischen DNA-Bildzytometrie	62
3.4.1	Auswertung nach Lage der DNA-Stammlinie	62
3.4.2	Auswertung nach Zahl der 9cEE.....	62
3.4.3	Kombinierte Auswertung nach Lage der DNA-Stammlinie und Zahl der 9cEE	63
3.5	Treffsicherheit der manuellen und der automatischen DNA-Bildzytometrie....	64
3.5.1	Treffsicherheit nach Lage der DNA-Stammlinie	64
3.5.2	Treffsicherheit nach Zahl der 9cEE.....	66
3.5.3	Treffsicherheit nach Kombination aus der Lage der DNA-Stammlinie und der Zahl der 9cEE.....	68
3.6	Diskrepante Fälle.....	70
4	Diskussion	71
4.1	Adjuvante Methoden in der Schilddrüsenzytologie	72
4.2	Methodischer und diagnostischer Vergleich der manuellen und der automatischen DNA-Bildzytometrie.....	76
4.2.1	Vergleichbarkeit der Kenngrößen der manuellen und der automatischen DNA-Bildzytometrie.....	78
4.2.2	Vergleichbarkeit der diagnostischen Interpretation der manuellen und der automatischen DNA-Bildzytometrie	80
4.3	Interpretation der 14 zwischen manueller und automatischer Messung diskrepanten Fälle	81
4.3.1	Sechs Fälle mit Unterschieden in der Interpretation des 9cEE	81
4.3.2	Fünf Fälle, in denen die in einer der beiden Messungen ermittelte aneuploide DNA-Stammlinie in der anderen Messung nicht bestätigt werden konnte	82
4.3.3	Drei Fälle mit hypodiploiden automatischen Messergebnissen, die man für eine klinische Interpretation nicht akzeptieren dürfte.....	84
4.3.4	Zusammenfassende Beurteilung.....	85

4.4	Diagnostische Treffsicherheit der manuellen und der automatischen DNA-Bildzytometrie in der Schilddrüsenzytologie.....	86
4.5	Vergleich der Ergebnisse mit der Literatur	88
5	Literaturverzeichnis	92
6	Abbildungsverzeichnis.....	105
7	Tabellenverzeichnis	107
	Danksagung	110

1 Einleitung

1.1 Epidemiologie knotiger Veränderungen der Schilddrüse

Knotige Veränderungen der Schilddrüse sind in Deutschland weit verbreitet und sind insbesondere in endemischen Jodmangelgebieten häufig (Reiners et al., 2004), (Führer et al., 2012) (Wienhold, 2013). Im Jodmonitoring des Kinder- und Jugendgesundheits-survey (KiGGS) wurden vor ca. 15 Jahren die Schilddrüsenvolumina aller Kinder und Jugendlichen ab 6 Jahren sonografisch ermittelt (Thamm et al., 2007). Zur Beurteilung der Jodversorgung der Bevölkerung wurde zusätzlich die Jodausscheidung im Urin gemessen. Die Jodurie lag mit 117 µg/l am unteren Ende der von der Weltgesundheitsorganisation (WHO) empfohlenen Menge (100–200 µg/l). In Anlehnung an die WHO-Kriterien zur Bewertung der Jodversorgung in der Bevölkerung (World Health Organization et al., 2007) entwickelt sich Deutschland inzwischen aber zu einem Land mit ausreichender Jodversorgung (Thamm et al., 2007). Dies ist v.a. auf die umfassende Bereitstellung und Verwendung von jodiertem Speisesalz (1 µg Jodid pro 1 g Salz) zurückzuführen.

Die häufigsten benignen Erkrankungen der Schilddrüse sind die euthyreote Struma diffusa et nodosa sowie Adenome und Entzündungen. In Deutschland weisen mehr als 30% der Erwachsenen eine Schilddrüsenvergrößerung oder einen Schilddrüsenknoten auf (Grußendorf, 2015). In der 2003 publizierten ersten Untersuchung der SHIP-Kohorte ("Study of Health in Pomerania", SHIP-0) wurden bei 35,9 % der 3941 Probanden mit bislang nicht diagnostizierter Schilddrüsenerkrankung eine Struma und bei 20,2% der Probanden Schilddrüsenknoten nachgewiesen (Völzke et al., 2003). In der 5-Jahres-Verlaufsuntersuchung der SHIP-Kohorte von 2008-2012 (SHIP-TREND) stieg die relative Häufigkeit von diagnostizierten Schilddrüsenerkrankungen von 7,6% in SHIP-0 auf 18,3% in SHIP-TREND (Völzke, 2012). Die relative Häufigkeit der Schilddrüsenmedikation stieg ebenfalls. Wurden in SHIP-0 nur 6,2% der Probanden medikamentös behandelt, so betrug der Anteil in SHIP-TREND 10,4%.

Frauen sind dreimal häufiger als Männer von Schilddrüsenerkrankungen betroffen (Melchert et al., 2002) (Melchert et al., 2005) (Reiners et al. 2004), wie auch in der SHIP-Kohorte erneut belegt werden konnte. Eine wesentliche Zunahme der Strumaprävalenz lässt sich außerdem mit zunehmendem Alter verzeichnen (Farahati et al., 2006). Während nur 5-8% der unter 25-Jährigen Schilddrüsenknoten aufweisen, sind es bei den über 55-Jährigen 32-45% (Siewert et al., 2012). Die Gesamtprävalenz der Struma liegt bei Frauen bei 3,5% in der jüngsten und bei 27,7% in der ältesten Altersgruppe. Bei den Männern steigt die Prävalenz von 4,5% auf 25,1% an (Reiners, Wegscheider, Schicha, et al., 2004). Je nach dem, welche Referenzwerte zur Definition einer Struma herangezogen werden, liegt die Prävalenz von vergrößerten Schilddrüsen bei 6-17-jährigen Kindern und Jugendlichen zwischen unter 5% und über 30% (Grußendorf, 2015).

Das follikuläre Schilddrüsenadenom ist neben der Knotenstruma eine weitere häufige Ursache einer Knotenbildung in der Schilddrüse und mit seinen Varianten der einzige gutartige Tumor des Schilddrüsenparenchyms (Thomas, 2006). Es kommt bevorzugt im mittleren Lebensalter vor, Frauen sind 7-mal häufiger betroffen als Männer (Krams & Frahm, 2019). Die Häufigkeit von Adenomen in Autopsiestudien wird mit 3-5% angegeben (Bisi et al., 1989), (Avetisian & Petrova, 1996).

Maligne Tumoren der Schilddrüse treten selten auf und machen weniger als 1% aller Raumforderungen der Schilddrüse aus (Reiners et al., 2008); allerdings ist ihre Inzidenz steigend und hat sich weltweit in den letzten 30–40 Jahren verdoppelt (Reiners et al., 2005). Laut den letztverfügbaren Krebsregisterdaten (RKI 2019) betrug die standardisierte Neuerkrankungsrate 2016 bei Frauen 11,1 und bei Männern 5,1 auf jeweils 100000; die Gesamtzahl an Neuerkrankungen betrug im gleichen Jahr 7780; für 2020 wurden 8800 Neuerkrankungen prognostiziert. Die Schilddrüsenkarzinome liegen damit in der Rangfolge des Anteils der häufigsten Tumorlokalisationen an allen Krebsneuerkrankungen 2016 (neueste zum Mai 2021 verfügbare Daten) auf Platz 19 (Männer) bzw. 12 (Frauen) (RKI 2019).

Mögliche Ursachen für die Zunahme könnten unter anderem die erhöhte Detektion auch kleiner, klinisch stummer Tumoren durch Zunahme an diagnostischen Aktivitäten wie Sonographie oder Feinnadelpunktion (FNP) (Leenhardt et al., 2004) und eine gestiegene Exposition gegenüber Umweltradioaktivität bei Kernwaffentests nach dem Zweiten Weltkrieg und der Tschernobyl-Katastrophe sein (Mürbeth et al., 2004) (Cotterill et al., 2001) (Magnanti et al., 2009). Auch der gravierende Atomunfall in Fukushima im Jahr 2012 dürfte in der Zukunft Auswirkungen zeigen.

Bezogen auf die Tumorentitäten ist vor allem eine Zunahme des papillären Karzinoms zu beobachten. Laut einer Studie (Radespiel-Tröger et al., 2014) über die Inzidenzzunahme von Schilddrüsenkarzinomen in Deutschland zwischen 1926 und 2009 machen papilläre Schilddrüsenkarzinome heute mit 69% bei Frauen und 59% bei Männern die häufigsten malignen Schilddrüsentumore in Deutschland aus. Der Anteil der follikulären Karzinome beträgt für Frauen 14%; für Männern 17%, gefolgt vom medullären Karzinom (M, 8%; F, 5%) sowie dem anaplastischen Karzinom (M, 5%; F, 3%).

Mit ca. 100.000 Schilddrüsenoperationen pro Jahr (Anonymus, 2013) weist Deutschland zum Beispiel im Vergleich gegenüber den USA mit 75.000 Eingriffen, aber etwa dreifacher Bevölkerungszahl (H. Li et al., 2011), eine deutlich höhere Zahl von Schilddrüsenoperationen auf. Schilddrüsenoperationen gehören somit zu den häufigsten Eingriffen in Deutschland (Pfannenstiel, 1993). Die meisten dieser Eingriffe dienen dem Ziel, eine knotige Schilddrüsenveränderung abzuklären und dabei insbesondere zur Abgrenzung von gutartigen Strumen und Adenomen gegenüber manifest malignen Tumoren. Die Zahl der Operationen benigner Strumen hat in Deutschland von 2005 nach 2013 zwar von 89000 auf 79000 pro Jahr abgenommen. (Bartsch et al., 2018). Jedoch ist die Malignitätsrate bei „diagnostischen Schilddrüsenoperationen“ wegen eines Schilddrüsenknotens im europäischen Vergleich

in Deutschland am geringsten. Das Verhältnis von Operationen mit maligner Tumordiagnose versus benigner Struma beträgt in Deutschland 1:15 (Dralle et al., 2004), in Italien 1:7 (Rosato et al., 2004) und in England und Skandinavien 1:5 (Scott-Coombes & Kinsman, 2007) (Bergenfels et al., 2008). Dass somit nur ein relativ kleiner Anteil der operierten knotigen Veränderungen tatsächlich maligne ist, führte zu der These, dass jährlich ca. 70000 Eingriffe an der Schilddrüse aus tumorbiologischer Sicht unnötig seien (Schmid und Tötsch, 2006). Eine andere Studie hat ergeben, dass die Zahl der Operationen um 87% gesenkt werden könnte, wenn alle verdächtigen Knoten vorher zytologisch überprüft werden würden (Carpi et al., 1996).

1.2 Klassifikation von Raumforderungen der Schilddrüse

An der Schilddrüse bzw. in der Schilddrüse sind eine Vielzahl von Raumforderungen möglich, darunter strumigene Veränderungen, Hämatome sowie entzündliche und postentzündliche Läsionen. Außerdem gibt es Hamartome und nicht selten auch intrathyroidale Lagen von Lymphknoten oder Nebenschilddrüsen. Am biologisch Wichtigsten sind schließlich die gutartigen und bösartigen epithelialen Tumoren, wohingegen nicht-epitheliale Tumoren keine große Rolle spielen.

Im Folgenden soll auf die wichtigsten gut- und bösartigen Erkrankungen mit Schwerpunkt auf das eigene Patientengut näher eingegangen werden.

1.2.1 Benigne Schilddrüsenerkrankungen

Die **Struma** beschreibt eine Vergrößerung der Schilddrüse. Diese kann diffus oder nodös sein bei euthyreoter, hypothyreoter oder hyperthyreoter Stoffwechselerkrankung. Die nicht-entzündliche, nicht-maligne Vergrößerung der Schilddrüse, die euthyreote Struma, kommt durch eine Vermehrung der Follikelzahl bei normaler Hormonproduktion zustande. Sie liegt in einem Endemiegebiet etwa 90 % aller Schilddrüsenerkrankungen zugrunde (Krohn et al., 2005). Die meisten Strumen sind asymptomatisch (Hegedüs et al., 2003).

Je nach Morphologie wird die Struma diffusa (gleichmäßig vergrößert) von der Struma nodosa (ein- oder mehrknotig) unterschieden. Von einer Struma diffusa spricht man, wenn das Schilddrüsenvolumen bei Frauen größer als 18 ml, bei Männern größer als 25 ml ist (Blank & Braun, 2007). Eine Struma nodosa liegt dann vor, wenn eine oder mehrere in der Schilddrüse gelegene Knoten tastbar oder bei entsprechender Größe sichtbar sind. Nicht selten weisen sie auch zystische Areale auf, ferner können sie stärkere regressive Veränderungen, aber auch eine Hyperplasie aufweisen. In Abhängigkeit von der Lage wird die eutope Struma von der dystopen Struma (z.B. Zungengrundstruma, abgrenzende laterale Halsstruma, intralaryngeale oder intratracheale bzw. auch retrosternale Struma) abgegrenzt (Hörmann, 2005).

Sind innerhalb eines Gebietes mehr als 10 % der Bevölkerung betroffen, spricht man von einer endemischen, ansonsten von einer sporadischen Struma. Als Auslöser für die sporadische Form gelten Hormonumstellungen, wie Pubertät, Klimakterium, Schwangerschaft oder strumigene Noxen (z.B. Lithium, Thiouracil, Perchlorat) (Zechmann & Haufe, 2012).

Das **follikuläre Adenom** ist der häufigste gutartige Schilddrüsentumor und mit seinen Unterformen die einzige gutartige epitheliale Tumorentität der Schilddrüse. Die Häufigkeit von Adenomen wird in Autopsiestudien mit 3-5% angegeben (Bisi et al., 1989), (Avetisian & Petrova, 1996).

Makroskopisch sind follikuläre Adenome meist solitäre umschriebene Knoten, die bis zu drei Zentimeter groß werden und eine weiche blasse Schnittfläche besitzen. Diese weist häufig Einblutungen, fibrotische Areale und Zysten auf. Solange die Adenome hormoninaktiv sind, bleiben sie häufig symptomlos. Hormonaktive Adenome (sogenannte dekompenzierte oder toxische Adenome) hingegen können eine Hyperthyreose auslösen (Krams & Frahm, 2019).

Histologisch erfordert die Diagnose eines follikulären Adenoms eine follikulär aufgebaute, gekapselte Neoplasie ohne papilläre Kerneigenschaften und ohne Invasion von Gefäßen oder Kapsel (Schmid & Farid, 2006). Sobald diese Kriterien nicht erfüllt sind, kommen differentialdiagnostisch verschiedene, z.T. erst kürzlich in der aktuellen WHO-Klassifikation von 2017 (Lloyd et al., 2017) vorgestellte Entitäten in Betracht, wie z.B. die Nicht-invasive follikuläre Neoplasie mit einem den papillären Tumoren ähnlichen Kerneigenschaften (NIFTP), der ein sehr niedrig-malignes Potential zugeschrieben wird. Ist eine eindeutige Invasion der Kapsel oder der Gefäße gesichert, kommt in Abhängigkeit von der Kernmorphologie differentialdiagnostisch ein follikuläres Schilddrüsenkarzinom oder die follikuläre Variante eines papillären Schilddrüsenkarzinoms in Betracht.

1.2.2 Maligne Schilddrüsenerkrankungen

Die histologische Klassifikation der Schilddrüsentumoren erfolgt nach der Einteilung der WHO (Lloyd et al., 2017). Demnach werden die Karzinome der Schilddrüse aufgrund ihrer Morphologie und des biologischen Verlaufs traditionell in vier (Haupt-) Entitäten unterteilt (DeLellis et al., 2004), wobei zusätzlich einige weitere Entitäten (gering differenziertes Schilddrüsenkarzinom (PDTC), onkozytäres Karzinom) nicht unerwähnt bleiben sollten:

- das follikuläre Karzinom (FTC, „follicular thyroid carcinoma“),
- das papilläre Karzinom (PTC, „papillary thyroid carcinoma“),
- das medulläre Karzinom (MTC, „medullary thyroid carcinoma“) und
- das anaplastische Karzinom (ATC, „anaplastic thyroid carcinoma“)

Follikuläre Karzinome der Schilddrüse sind diejenigen differenzierten Tumore der Schilddrüse, die follikulär aufgebaut sind, von Thyreozyten ausgehen und nicht die zytologischen (Milchglaskerne) oder histologischen (Papillen) Kriterien papillärer Karzinome erfüllen (DeLellis et al., 2004) (Lloyd et al., 2017). Morphologisch ist das follikuläre Schilddrüsenkarzinom meist von einer bindegewebigen Kapsel umgeben. Als entscheidendes Malignitätskriterium gelten Kapseldurchbrüche mit Infiltration in das umgebende Gewebe sowie Gefäßeinbrüche (Schmid & Farid, 2006). Je nach Ausmaß der Invasion unterscheidet man gemäß WHO-Klassifikation drei Formen des follikulären Karzinoms, die von prognostischer Relevanz sind (Lloyd et al., 2017).

- Minimal-invasives (gekapseltes) follikuläres Karzinom (MIFTC): Diese Tumoren stellen etwa zwei Drittel der follikulären Karzinome dar (Asari, 2009) und haben eine deutlich bessere Prognose als grob invasive (Dralle & Machens, 2008).
- Gekapseltes, angioinvasives follikuläres Karzinom: Diese Karzinome zeigen ein limitiertes (<4 Gefäße) oder ein ausgedehntes (≥ 4 Gefäße) angioinvasives Wachstum.
- Grob-invasives Karzinom: Diese oft großen Karzinome weisen zahlreiche Gefäßeinbrüche und Kapseldurchbrüche (Schmid, 2010) sowie oft eine ausgedehnte Infiltration des Schilddrüsenorgans mit oder ohne extrathyreoidale Ausbreitung auf.

Das follikuläre Karzinom metastasiert vorzugsweise hämatogen. Die Prognose ist insgesamt deutlich schlechter als beim papillären Karzinom; die 5- und die 10-Jahresüberlebensrate betragen aber aktuell 84,9% bzw. 72,9% in der insgesamt fast 8000 Fälle umfassenden Tumorkohorte des Tumorzentrums München aus den Jahren 1998-2020 (Tumorregister München 2021), in der 782 Fälle von follikulären Schilddrüsenkarzinomen nachverfolgt wurden. Die Prognose hängt von verschiedenen Faktoren ab: Alter der Patienten, Größe und Staging des Tumors, Vollständigkeit der Operation und Reaktionsfähigkeit gegenüber radioaktivem Jod (Sobrinho-Simões et al., 2011).

Das **papilläre Schilddrüsenkarzinom** (PTC) zeigt morphologisch vor allem eine papilläre Architektur mit neoplastischen Follikeln und soliden Tumorzellaggregaten; papilläre Strukturen kommen aber auch in hyperplastischen Knoten oder Adenomen vor (Rothmund, 2007). Die Zellkerne sind meist auffällig chromatinarm (Milchglaskerne) und überlagern sich dachziegelartig (Bramley & Harrison, 1996), (Pimpl et al., 2013); diese Kriterien sind insgesamt für die Abgrenzung papillärer gegenüber anderen Karzinomen entscheidend. Darüber hinaus besteht oft eine größere Variabilität der Kerne mit Entrundungen und Einfaltungen der Kernmembran, sowie gelegentlich auch eine Hyperchromasie. Eine weitere diagnostisch hilfreiche Besonderheit des papillären Schilddrüsenkarzinoms ist das Auftreten von Psammomkörperchen, konzentrischen lamellierten verkalkten Strukturen mit einer Größe von etwa 50 bis 70 μm (Cai et al., 2015).

Papilläre Strukturen sind häufig, fehlen jedoch bei der follikulären Variante. Die follikuläre Variante des papillären Schilddrüsenkarzinoms zeigt architektonisch ein

vorwiegend follikuläres Wachstum. Die Follikel sind nicht wie im normalen Schilddrüsengewebe von flachem Epithel ausgekleidet, sondern von Zellen umgeben, die der typischen Morphologie des bereits beschriebenen PTC entsprechen (Salajegheh et al., 2008). Die WHO klassifiziert das papilläre Schilddrüsenkarzinom neben der follikulären Variante des Weiteren in die entkapselte, die Zylinderepithel- (Columnar cell-) sowie die onkozytäre Variante des PTC. Eine weitere Gruppe stellen die papillären Mikrokarzinome dar.

Das PTC metastasiert überwiegend lokal lymphogen, Fernmetastasen sind eher selten (Al-Brahim & Asa, 2006). Insgesamt betragen in der Münchener Kohorte von 6309 Fällen die 5- und die 10-Jahresüberlebensraten 97,3% und 93,0% (Tumorregister München 2021) und werden auch sonst mit > 90% angegeben (Nikiforov et al., 2016). Letztendlich besteht, wenn man nicht das absolute, sondern das relative Überleben zugrunde legt, eine normale Lebenserwartung mit Überlebensraten von 100,4% nach 5 Jahren und 100,2% nach 10 Jahren (Tumorregister München 2021).

Das **medulläre Schilddrüsenkarzinom** ist selten und macht etwa 4%–10% aller Schilddrüsenkarzinome aus (Lörincz et al., 2016). Es handelt sich um einen malignen neuroendokrinen Tumor der parafollikulären C-Zellen, der durch langsames Wachstum und sehr frühe lymphogene Metastasierung charakterisiert ist (Karges, 2010). Das histologische Bild kann stark variieren. Der Tumor kann solide, follikuläre, oder seltener (pseudo-) papilläre, onkozytäre oder kleinzellige Areale enthalten. Aufgrund der frühen Metastasierungstendenz und des Nicht-Ansprechens auf eine Radiojodtherapie ist die Tumorprognose insbesondere bei fortgeschrittenen Fällen ungünstiger. Sie ist abhängig von der Tumorausbreitung zum Zeitpunkt der Diagnose. Während beim ausschließlich auf die Schilddrüse beschränkten MTC das Langzeitüberleben >95% beträgt, verschlechtert sich die Prognose bei Metastasierung in die Lymphknoten am Hals oder das obere Mediastinum deutlich (Tiedje et al., 2015). Die entsprechenden Werte der 360 Fälle umfassenden Münchener Kohorte für das 5- und das 10-Jahres-Überleben liegen bei 91,7% bzw. bei 83,3% (Tumorregister München 2021).

Anaplastische (undifferenzierte) Karzinome treten meist im hohen Alter auf, Frauen sind doppelt so häufig betroffen wie Männer (Besic et al., 2005).

Anaplastische Karzinome infiltrieren schnell in die Halsweichteile und metastasieren frühzeitig lymphogen und hämatogen in Lunge und Leber (Mann, 2002). In Deutschland macht es ca. 4 % der Schilddrüsenkarzinome aus (Krebs - Schilddrüsenkrebs, 2016). Morphologisch ist es gekennzeichnet durch hochgradig atypische, meist einzeln gelegene Tumorzellen neben Zelldetritus und gelegentlich auch Entzündungszellen (Tötsch et al., 2005). Die Prognose ist im Wesentlichen schlecht mit einem nur geringen Langzeitüberleben (Führer et al., 2019). Für die 201 Münchener Fälle wird ein 3-Jahres-Überleben von 13,2% angegeben (Tumorregister München 2021).

1.3 Präoperative Diagnostik von Schilddrüsenknoten

Wichtig für einen rationalen Umgang mit knotigen Schilddrüsenveränderungen ist eine zielgerichtete Diagnostik, die mit dem Ziel erfolgt, eine therapeutische Entscheidung auf einer möglichst gesicherten Basis von Erkenntnissen herbeizuführen. Dabei stehen zum einen Anamnese und klinische Untersuchung zur Verfügung, zum anderen die bildgebende Diagnostik und die Anwendung morphologischer Untersuchungsverfahren der Zytologie und in begrenztem Maße auch der Histologie.

1.3.1 Anamnese und klinischer Untersuchungsbefund

Bei der Anamnese ist nach Art, Dauer und möglicher Progression einer Knotensymptomatik (u. a. Hyperthyreosezeichen, klinische Malignitätszeichen, mechanische Beeinträchtigung) und einer zurückliegenden Bestrahlung der Kopf- und Hals-Region (erhöhtes Risiko für Schilddrüsenmalignom) zu fragen. Bei der Familienanamnese sind Hinweise für eine hereditäre Schilddrüsenerkrankung, z. B. multiple endokrine Neoplasie Typ 2, das Cowden-Syndrom, familiäre Schilddrüsenautonomien oder familiäre Struma von Bedeutung (Führer & Schmid, 2010). Anamnestische und klinische Faktoren, die für ein erhöhtes Malignitätsrisiko sprechen, sind ein Patientenalter < 14 Jahre oder > 70 Jahre, männliches Geschlecht, wachsender und / oder nicht verschieblicher Knoten, harte Knotenkonsistenz, vergrößerte Halslymphknoten, Heiserkeit, Dysphagie und Husten (Gharib et al., 2016).

1.3.2 Sonographie

Die Sonographie ist unentbehrlicher Bestandteil der Basisdiagnostik der Schilddrüse. Mit Hilfe des Ultraschalls lassen sich die Echostruktur (homogen oder inhomogen) und die Echogenität (echogleich, echofrei, echoarm oder echoreich) der Schilddrüse beurteilen. Beim Echoverhalten vergleicht man die umliegende Halsmuskulatur mit der Schilddrüse. Normalerweise grenzt sich die Schilddrüse im Vergleich zur Halsmuskulatur deutlich heller ab. Man bezeichnet das Verhalten als echodicht und im Vergleich zur normalen Schilddrüse echonormal. Erscheint eine Struktur ähnlich wie die Halsmuskulatur und somit dunkler als das gesunde Schilddrüsengewebe, spricht man von echoarm. Echonormale und echoreiche Knoten, teils mit echoarmem Randsaum (Zone vermehrter Durchblutung) und echofreien Anteilen, sind häufig nachweisbar und entsprechen meist gutartigen Veränderungen. Als echofrei erscheinen beispielsweise Zysten.

Ferner können im Ultraschall benachbarte Strukturen, unklare Raumforderungen (diffuse, uni- oder multifokale Läsionen) sowie Lymphknoten im Halsbereich, Kalzifikation, die Knotenanzahl, die Größe der Schilddrüse und der Schilddrüsenknoten und die Durchblutung erfasst werden (Frates et al., 2005).

Ziel der Ultraschalluntersuchung ist die Stratifizierung des Malignitätsrisikos des Schilddrüsenknotens (Georgopoulos et al., 2003). Als Malignitätshinweis gelten vergrößerte und verdächtige Lymphknoten und das invasive Wachstum eines Knotens. Weitere mit Malignität assoziierte sonographische Charakteristika von Schilddrüsenknoten sind u.a. Mikrokalzifikation, Hypoechogenität, Mikrolobulierung oder unregelmäßige Begrenzung, fehlender Halo, solide Knoten und intranoduläre Vaskularisierung (Lehnert, 2014). Solche im TI-RADS (thyroid imaging reporting and data system) zusammengefassten Kriterien (Kwak et al., 2011) sind jedoch variabel und untersucherabhängig (Wienke et al., 2003), (Kwak et al., 2011).

Eine Weiterentwicklung der Ultraschalldiagnostik ist die farbkodierte Doppler-sonographie. Sie ermöglicht eine qualitative Beurteilung der Gefäßarchitektonik sowie die Quantifizierung der Gewebevaskularisation. Tumorwachstum ist mit Hypervaskularisation im betroffenen Gebiet assoziiert. Eine rein perinodale Durchblutung im Doppler spricht für eine hohe Benignitätswahrscheinlichkeit. Mit Zunahme der zentralen Vaskularisation eines Knotens steigt auch das Risiko für ein Karzinom (Radl et al., 2012).

1.3.3 Laborparameter

Für die Schilddrüsendiagnostik sind verschiedene Parameter von Bedeutung. Durch Bestimmung von Thyreotropin (TSH) im Serum lassen sich Schilddrüsenfunktionsstörungen erkennen, die im Rahmen einer uni- oder multifokalen funktionellen Autonomie auftreten können (Pacini et al., 2006) (Cooper et al., 2009). Freies T3 und T4 (fT3, fT4) eignen sich zur Abklärung einer Hyperthyreose, wenn TSH supprimiert ist (Bruhn et al., 2011). Bei Verdacht auf eine Autoimmunerkrankung werden Schilddrüsenautoantikörper (Thyreoglobulin-Ak, Ak gegen thyreoidale Peroxidase, TSH- Rezeptorautoantikörper) bestimmt.

Serumthyreoglobulin spielt für die Primärdiagnostik keine Rolle, da der Wert auch bei gutartigen Erkrankungen erhöht sein kann (Guerrero et al., 2010). Es findet Verwendung bei der Rezidivdiagnostik differenzierter Schilddrüsenkarzinome nach totaler Thyreoidektomie (Kloos, 2010). Die Bestimmung des Serumcalcitonins ist bei der Erkennung des medullären Karzinoms von Bedeutung (Karges et al., 2004).

1.3.4 Bildgebende Verfahren

Die Szintigraphie als ergänzendes Verfahren zur sonographischen Untersuchung der Schilddrüse ermöglicht, neben der TSH-Messung, morphologische Veränderungen hinsichtlich ihres Funktionszustandes zu charakterisieren. Das Verfahren wird meist mit ^{99m}Tc -Pertechnetat durchgeführt und liefert Hinweise darauf, ob Knoten szintigraphisch kalt oder heiß sind und wie groß die funktionelle Aktivität der Schilddrüse oder der einzelnen Schilddrüsenknoten ist (Haugen et al., 2016).

Ein mittels ^{99m}Tc -Pertechnetat szintigraphierter hyperfunktioneller Knoten schließt ein Schilddrüsenkarzinom mit großer Wahrscheinlichkeit aus, sodass eine weitere Abklärung zur Dignitätsabschätzung, z.B. durch eine zytologische Untersuchung, nicht erforderlich ist (Rosen & Stone, 2006). Kalte Knoten weisen hingegen ein erhöhtes Malignitätsrisiko auf und sollten daher der Feinnadelbiopsie zugeführt werden (Schmidt, 2016).

Andere bildgebende Verfahren wie die Computertomographie (CT) und Magnetresonanztomographie (MRT) haben in der Primärdiagnostik von Raumforderungen der Schilddrüse keine wesentliche Bedeutung. Die Indikation zu CT oder MRT ist bei Verdacht auf ausgedehnte lokoregionäre oder intrathorakale Ausbreitung und Metastasierung gegeben. Schilddrüsenläsionen, die mittels CT oder MRT als bösartig eingestuft werden, sollten zusätzlich mit einer sonographischen Untersuchung abgeklärt werden (Shetty et al., 2006).

1.3.5 Zytologische Untersuchung

Die Schilddrüsenzytologie fand in der Diagnostik erstmals Anwendung in Skandinavien durch Söderström (Söderström, 1952). Mittlerweile gilt sie als eines der Standardverfahren zur Beurteilung, ob eine chirurgische Entfernung eines erkannten Knotens notwendig ist oder nicht (Langer et al., 2012).

Die Zellen werden dabei durch eine Feinnadelpunktion gewonnen. Die übliche Punktion wird mit einer Nadelstärke von 25-27 Gauge durchgeführt, die ein minimales Risiko einer Blutung bei suffizienter Materialgewinnung gewährleisten soll. Die Verwendung von Nadeln mit einem größeren Binnendurchmesser verbessert nicht den Ertrag an zytologisch auswertbarem Material, sondern geht mit einem gering höheren Blutungsrisiko und einer höheren Rate an hämorrhagischen Blutausstrichen ohne Thyreozyten einher (İnci et al., 2013). Die Aspiration erfolgt entweder per Hand durch Sogwirkung mit einer 10- bis 20-ml-Spritze oder durch die Verwendung einer Aspirationshilfe wie z.B. einem Cameco-Griff. Die Nadel wird dabei fächerförmig im Knoten vor- und zurückgeschoben. Nachfolgend wird das aufgenommene Material auf einem oder mehreren Objektträgern ausgestrichen. Bei luftgetrockneten Präparaten erfolgt die Färbung gemäß den SOPs (*Standard Operating Procedures*) mit einer hämatologischen Färbung, z.B. nach May-Grünwald-Giemsa, nach alkoholischer Fixation mit einer Färbung nach Papanicolaou oder einer Hämatoxylin-Eosin-Färbung. Als Alternative zu Direktausstrichen kann das Punktat auch in ein Fixationsmedium eingespritzt und weiter flüssigkeitsbasiert prozessiert werden (Dünnschichtzytologie).

Die Durchführung, wie z.B. die Anzahl der Punktionen pro Knoten, wird unterschiedlich gehandhabt. Entscheidend ist in jedem Fall die Gewinnung einer für den Schilddrüsenknoten repräsentativen und für die Beurteilung ausreichenden Zellmenge (Karger et al., 2008).

Indikationen zur Schilddrüsenzytologie

Die Schilddrüsenzytologie wird als Referenzmethode zur Beurteilung fokaler Läsionen der Schilddrüse angesehen. Dabei geht es zum einen um die Bestimmung des Malignitätsrisikos von Schilddrüsenknoten und zum anderen- bei nachgewiesenem Malignom- der Bestimmung des Tumortyps (Lundgren et al., 2008) (Schmid & Reiners, 2011).

In Struma-Endemiegebieten, in denen knotige Schilddrüsenveränderung à priori häufig sind, kann nicht jede Läsion punktiert werden; hier würde die Punktion jeden Knotens selbst ab 1 cm Größe zu einer unnötig erhöhten Operationsfrequenz führen (Ting et al., 2015). Eine Empfehlung zur Punktion ergibt sich bei folgenden Konstellationen (Cooper et al., 2009) (Gharib et al., 2010) (Perros et al., 2014):

- Patienten mit klinischen Zeichen eines Schilddrüsenkarzinoms
- Knoten >1 cm mit mindestens zwei sonographischen Malignitätskriterien
- Knoten jeder Größe mit extrakapsulärem Wachstum oder unklaren zervikalen Lymphknoten
- Knoten jeder Größe bei Patienten mit einer Bestrahlung im Halsbereich in der Vorgeschichte
- bei mehr als zwei erstgradigen Verwandten mit differenziertem Schilddrüsenkarzinom
- bei medullärem Schilddrüsenkarzinom oder multipler endokriner Neoplasie (MEN) Typ 2
- bei erhöhten Calcitonin-Werten.

Im Nicht-Strumaendemiegebiet geht es hingegen darum, durch eine Feinnadelpunktion der Schilddrüse allgemein die Genese der dort selteneren knotigen Läsionen zu klären. Die Sensitivität der Malignomerkenung soll hier möglichst hoch sein. Das heißt, bereits bei entsprechendem zytologischem Verdacht wäre eine Operation indiziert. Eine Empfehlung aus Deutschland findet sich aus der Arbeitsgruppe um Schmid (Ting et al. 2015): Nach szintigrafischem Ausschluss einer Autonomie und bei auffälliger Klinik bzw. auffälligem Ultraschallbefund soll eine Feinnadelpunktion angestrebt werden.

Diagnostische Kriterien der Schilddrüsenzytologie

Bei der zytologischen Beurteilung zytologischer Proben werden bei den allermeisten Materialien die gleichen oder zumindest vergleichbare diagnostische Kriterien angelegt (zur Übersicht siehe z.B. Koss und Melamed 2005 bzw. Bubendorf et al. 2011a). Insbesondere geht es in der Regel um Veränderungen des Kernbildes von Einzelzellen (Hyperchromasie, Chromatinvergrößerung, Beurteilung der Kern-Plasma-Relation etc.) bzw. von einer Zellpopulation (Kerngrößenvariabilität, Nekroseneigung, Zelldissoziation etc.) im Vergleich zu einer Referenz von normalen Zellen des gleichen Zelltyps.

Außerdem können auch andere Zellarten, die üblicherweise in den Proben zu erwarten sind, zum Vergleich herangezogen werden; hierzu können andere Epithelarten gehören, aber auch Entzündungszellen oder Makrophagen.

Für die Schilddrüse sind diese Konzepte allerdings nur bedingt anwendbar. Gerade die mit großem Abstand häufigste Fragestellung, nämlich die nach dem biologischen Charakter einer follikulären knotigen Läsion, entzieht sich der Beurteilung nach den üblichen Kriterien der Zytologie. Dies liegt daran, dass follikulär differenzierte Schilddrüsenepithelien sowohl im Normalzustand, als auch bei Aktivierung im Rahmen von endokrinen Störungen und auch bei follikulären Tumoren meistens keine klaren Kernatypien aufweisen, sondern in ihren Kernen rundlich und in ihrem Chromatin überwiegend fein bleiben. Entsprechend weisen auch zytologische Präparate aus follikulär aufgebauten knotigen Läsionen keine wesentlichen Kernatypien auf, unabhängig davon, ob sie aus einer knotigen Kolloidstruma, einem follikulären Adenom oder einem follikulären Karzinom stammen.

Wenn somit die Zytologie bei diesen differentialdiagnostischen Fragestellungen nicht auf das klassische Repertoire an Kriterien zurückgreifen kann, so begründet sich hieraus, weswegen in der Schilddrüsenzytologie follikulärer Veränderungen andere Kategorisierungen der zytologischen Zellbilder vorgenommen werden müssen als in der übrigen Zytologie.

Stattdessen werden Hilfsgrößen wie z.B. der Epithelzellgehalt der Probe ("Zellularität") oder die Kerngrößenvariabilität ("Anisokaryose") zur Diagnostik herangezogen (zur Übersicht siehe Droese 1995, Bubendorf et al. 2011b), die aber nur relativer Natur sind. Letztendlich muss man sie sich im Wesentlichen korreliert mit der Größe der Follikel in der punktierten Läsion vorstellen. Makrofollikuläre Läsionen werden entsprechend weniger Zellen pro Aspirationsvolumen aufweisen als mikrofollikuläre. Da es andererseits sowohl bei Strumen, als auch bei follikulär differenzierten Tumoren makrofollikuläre und mikrofollikuläre Formen gibt, lässt sich ableiten, wie relativ das Kriterium der Zellularität insgesamt ist. Anisokaryose wiederum ist ebenfalls nur eine relative Größe, da sie vor allem den endokrinen Zustand der Läsion widerspiegelt. Endokrin aktivere, szintigraphisch warme Veränderungen werden vom Prinzip her zu größeren Schwankungen der Kerndurchmesser neigen als endokrin weniger aktive, wie man von typischen Zellbildern bei Hyperthyreose weiß. Andererseits gibt es auch bei Tumoren endokrin recht aktive, szintigraphisch warme oder heiße Formen, so dass auch dieses Kriterium relativer Natur ist.

Diese Relativierungen dürfen aber nicht dazu führen, dass der Eindruck entsteht, die Durchführung einer schilddrüsenzytologischen Befundung entspräche im Grunde einer unverbindlichen Abschätzung. Auf die Aussagekraft der Methode wurde vielfach hingewiesen, und schon Droese, der als Pionier der Schilddrüsenzytologie in Deutschland gelten darf und in der Zytopathologischen Abteilung des Universitätsklinikums

Göttingen diese Methode etabliert hatte, stellte bereits vor gut 40 Jahren heraus (Droese, 1979), dass die Schilddrüsenzytologie eine wichtige diagnostische Screeningmethode repräsentiert, mit deren Hilfe bei knotigen Veränderungen wesentliche Erkenntnisse darüber gewonnen werden können, ob diese operativ entfernt werden sollten oder belassen werden können. Im Besonderen wies er dabei auf die Notwendigkeit hin, eine Korrelation der zytologischen Befunde mit klinisch-bildgebenden Daten herbeizuführen.

Für andere, nicht follikuläre Veränderungen sind hingegen zytologische Diagnosen unter Anwendung der klassischen Kriterien der Zytologie möglich (Bubendorf et al. 2011b). So kann man ein anaplastisches Karzinom an seinen ausgeprägten Atypien erkennen, und auch für papilläre und medulläre Karzinome sowie für manche entzündliche Erkrankungen (Hashimoto- und de Quervain-Thyreoiditis) gibt es konventionelle zytologische Kriterien, die eine zytologische Diagnostik prinzipiell erlauben.

Kategorisierung zytologischer Schilddrüsenbefunde

Die Besonderheiten in der Schilddrüsenzytologie bilden sich auch darin ab, dass die Nomenklatur der Befunde von der üblichen abweicht. Statt der sonst gebräuchlichen viergliedrigen Kategorisierung als negativ, zweifelhaft, dringend verdächtig oder positiv, die von der Deutschen Gesellschaft für Zytologie (DGZ) und der Deutschen Gesellschaft für Pathologie (DGP) gemeinsam vorgeschlagen wurde (Böcking 1998b), haben sich bei der Schilddrüse abweichende Konzeptionen ergeben, die den organspezifischen Problemen Rechnung tragen.

Die Europäische Schilddrüsengesellschaft empfiehlt für die zytologische Auswertung folgende Einteilung (Gharib et al. 2010, Gharib et al. 2016, Feldkamp et al., 2016):

- I: unzureichendes Untersuchungsmaterial
- II: kein Nachweis von Tumorzellen, zytologisch benigne
- III: follikuläre Neoplasie unklarer Dignität
- IV: Malignomverdacht
- V: Zytologisch maligne

Diese Klassifikation, die noch ihre Herkunft aus der Pap-Klassifikation in der gynäkologischen Zytologie erkennen lässt, betrachtet neben einer eigenen Kategorie für die zytologisch nicht als gut- oder bösartig einzuordnenden follikulären Neoplasien je eine Kategorie für malignitätsverdächtige und sicher maligne Läsionen. Ähnlich wird auch am Funktionsbereich Cytopathologie in Düsseldorf verfahren (siehe Kapitel 2.3).

International relativ gebräuchlich ist ferner in den USA entwickelte Bethesda-Klassifikation für die Schilddrüsenzytologie (Schäffer et al., 2012) (Cibas & Ali, 2017).

Diese ist eine internationale Einteilung in sechs Befundgruppen, denen jeweils ein in Studien ermitteltes Malignitätsrisiko zugeordnet wird.

- I: nicht diagnostisch / vereinbar mit Zysteninhalt
- II: benigne / vereinbar mit einem follikulären Knoten
- III: Atypien unklarer Signifikanz
- IV: (V. a.) follikuläre Neoplasie
- V: Malignomverdacht
- VI: Zytologisch maligne

Letzten Endes erkennt man in beiden Klassifikationen das im Resultat sich nur wenig unterscheidende Bemühen, die verschiedenen optionalen Enddiagnosen follikulärer Läsionen in gut nutzbare Kategorien umzusetzen. Außerdem lässt sich ablesen, wie groß die Bandbreite möglicher Zuordnungen follikulärer Läsionen ist.

Diagnostische Feinnadelpunktate sollten gemäß den Empfehlungen der europäischen und US-amerikanischen Schilddrüsengesellschaften wie folgt therapeutisch interpretiert werden:

- Benigne Knoten (ca. 70%) erfordern keine weitere Diagnostik.
- Bei malignen Knoten (< 5%) ist ein primärer chirurgischer Eingriff zur operativen Tumorentfernung indiziert.
- Bei suspektem Knoten / follikulärer Neoplasie (ca. 15–20%) ist eine histologische Abklärung erforderlich, da die zu erwartende Malignomrate bei ca. 20% liegt (Frates et al., 2005) (Pacini et al., 2006) (Cooper et al., 2006).

Zusammenfassend sollte bei technisch inadäquaten Präparaten die Punktion wiederholt werden, gutartige Befunde sollten klinisch kontrolliert werden, sofern nicht besondere Operationsgründe, wie z.B. eine Dyspnoe durch tracheale Kompression einer Struma vorliegen. Bei allen nicht sicher gutartigen Befunden wäre die operative Abklärung zu empfehlen.

1.4 DNA-Bildzytometrie

Die DNA-Bildzytometrie stellt eine Kombination aus digitaler morphologischer Klassifizierung von Kernen verschiedener Zelltypen und quantitativer Messung des DNA-Gehaltes in den als morphologisch auffällig bewerteten Zellen dar (Böcking et al., 2015). Ihr Zweck ist im Wesentlichen die Erkennung bösartiger Tumorzellen (diagnostische DNA-Bildzytometrie), bei anderen Anwendungen aber auch die Bestimmung des Malignitätsgrades eines bekannten Tumors (prognostische DNA-Bildzytometrie) (Böcking et al., 2015). Es handelt sich um eine international im Rahmen von mehreren Konsensus-Konferenzen der ESACP ("European Society of Analytical and Cellular Pathology") standardisierte Methode (s. unten).

Nach quantitativer Färbung der Zellkern-DNA mithilfe der Feulgen-Färbung (Feulgen & Rossenbeck, 1924) können konventionelle Präparate der Punktionszytologie und der Exfoliativzytologie sowie Gewebe aus Paraffinblöcken (nach enzymatischer Zellvereinzelung) mit Hilfe eines TV-Bildanalysesystems einer DNA-zytometrischen Messung unterzogen werden.

Die Einsatzfähigkeit der DNA-Bildzytometrie in der Diagnostik und Gradierung von Tumoren basiert auf der Tatsache, dass die oft aneuploiden Tumorzellen einen anderen DNA-Gehalt aufweisen als normale bzw. reaktiv-hyperplastische Zellen. Mit Ausnahme der Keimzellen enthalten nicht-neoplastische Zellen 2×23 Chromosomen, d.h. einen doppelten Chromosomensatz von 46 Chromosomen, dessen DNA-Gehalt als $2c$ gesetzt und als diploid bezeichnet wird. Diejenigen Zellen, die sich in Proliferation (S-Phase) befinden, durchlaufen alle DNA-Gehalte bis $4c$, um sich in der G_2/M -Phase in zwei Zellen à $2c$ zu teilen. In manchen Geweben kommt es aber zu regelhaften Vervielfachungen des Chromosomensatzes, entsprechend den ganzzahligen Potenzen des $2c$ -Wertes, also $4c$, $8c$, $16c$, $32c$ etc. Dieser Vorgang ist Folge davon, dass sich bei $4c$ die Zelle nicht teilt, sondern einen oder mehrere weitere Zellzyklen durchläuft. Dieses Phänomen wird als "euploide Polyploidisierung" bezeichnet und findet sich beispielsweise in Epithelien der Schilddrüse, der Samenblasen, der Leber und der ableitenden Harnwege sowie in Mesothelien und Herzmuskelzellen. Die euploide Polyploidisierung ist bei diesen Geweben auf der einen Seite zwar physiologisch, kann aber durch Erkrankungen verstärkt werden, z.B. am Herzmuskel durch Hypertrophie bei einem angeborenen Vitium, oder auch durch erworbene Erkrankungen ausgelöst werden (zur Übersicht siehe Biesterfeld et al., 1994).

Die meisten Tumoren zeigen numerische und / oder strukturelle Chromosomenaberrationen (= chromosomale Aneuploidie) (Böcking, 2002). Das quantitative zytometrische Äquivalent chromosomaler Aneuploidie ist die DNA-Aneuploidie (Böcking, 1998a). Der Nachweis chromosomaler Aneuploidie bzw. numerischer Aneuploidie gilt international als Marker für das Vorliegen von neoplastischen Zellen (Nadjari et al., 1999) (Haroske et al., 2001a) (Böcking et al., 2014).

Messgröße bei der DNA-Bildzytometrie ist der DNA-Gehalt der Zellkerne, um den "Nettoeffekt" chromosomaler Aberrationen auf den Zellkern-DNA-Gehalt als diagnostischen Marker zu nutzen (Böcking 1998a). Der DNA-Gehalt selbst kann dabei durch die Zytometrie nicht direkt bestimmt werden. Die integrierte optische Dichte (IOD) des Kerns ist nach quantitativer Färbung nach Feulgen das zytometrische Äquivalent seines DNA-Gehaltes (Böcking et al., 1994). Die Messung des nukleären DNA-Gehaltes erfordert eine Skalierung der IOD-Werte durch einen Vergleich mit den Werten von Zellen mit bekanntem DNA-Gehalt. Sie werden Referenzzellen oder Eichzellen genannt und haben definitionsgemäß einen DNA-Gehalt von $2c$. Als Referenzzellen werden im zu messenden Präparat enthaltene normale gutartige Zellen wie z.B. Entzündungszellen, Fibroblasten oder normale Epithelzellen verwendet.

Die Häufigkeitsverteilung der IOD der Zellkerne wird in einem Histogramm aufgetragen und in der Einheit „c“ wiedergegeben, wobei 1c der Hälfte des Zellkern-DNA-Gehalts von Zellen einer normalen, nicht pathologischen, diploiden Population in der G₀/G₁-Phase des Zellzyklus entspricht (Böcking, 2002). Der modale Wert eines Histogramm-Peaks repräsentiert den häufigsten Wert (Gipfel) innerhalb dieses Peaks (Böcking & Nguyen, 2004).

Eine Population normaler, nicht proliferierender Zellen zeigt in einem DNA-Histogramm einen ersten Häufigkeitsgipfel bei 2c (entsprechend der G₀/G₁-Phase) und kann einen zweiten bei 4c aufweisen (entsprechend der G₂/M-Phase im Zellzyklus) (Haroske et al., 2001). Ferner können Einzelwerte zwischen 2c und 4c als Ausdruck von Zellen in der S-Phase vorkommen. Solche DNA-Verteilungen werden als "diploid" bezeichnet. Als aneuploid bezeichnet man in der DNA-Bildzytometrie DNA-Verteilungen von Zellpopulationen, die sich statistisch signifikant von denen normaler Zellen (ruhend, proliferierend, polyploid) unterscheiden (Böcking et al., 1994).

Das Ziel der DNA-Zytometrie ist somit das Erkennen von nicht normalen (aneuploiden) DNA-Stammlinien, die statistisch signifikant von den Lagen normaler diploider Stammlinien abweichen, ferner auch die Identifizierung von Zellen mit besonders hohem DNA-Gehalt, z.B. von > 9c. Physiologische Faktoren wie euploide Polyploidisierung oder Hyperregeneration ebenso wie unphysiologische Faktoren wie Apoptose, Virusinfektionen, Strahlen- und Zytostatikatherapie oder Vitamin B12-Mangel nehmen Einfluss auf den DNA-Gehalt von Einzelzellen, so dass diese Effekte bei der Interpretation von DNA-Histogrammen berücksichtigt werden müssen (Böcking, 1995).

Die DNA-Bildzytometrie wird am Funktionsbereich Cytopathologie der Universität Düsseldorf als ergänzende diagnostische Methode zur konventionellen Zytologie eingesetzt. Die Methode gilt dabei als sensitiv und reproduzierbar. Es werden zwar nur wenige hundert Zellen gemessen, kleinere Abweichungen im DNA-Gehalt können dabei aber problemlos erfasst werden. Außerdem ist es in Abgrenzung zum Leistungsspektrum der Durchflusszytometrie vorteilhaft, dass durch die visuelle Auswahlmöglichkeit der Zellen ein unerwünschtes Messen von unauffälligem Normalgewebe oder Zellartefakten vermieden werden kann. Darüber hinaus ermöglicht die interaktive Detektion morphologisch auffälliger Strukturen den Nachweis kleiner aneuploider Tumorzellpopulationen. Die Methode ist durch die subjektive Auswahl der Analysezellen durch den Untersucher aber auch abhängig von dessen Erfahrung. Ein heute nicht mehr ins Gewicht fallender weiterer Nachteil war, dass nicht von Anbeginn der klinischen Anwendung bereits Standards entwickelt waren, wie sie heute allgemein anerkannt sind (Haroske et al., 2001).

Auf die Erfahrungen in der wissenschaftlichen Literatur zur DNA-Zytometrie speziell an der Schilddrüse wird in Kapitel 4.4 ausführlicher eingegangen.

Hier soll nur darauf hingewiesen werden, dass die Schilddrüsenepithelien ähnlich wie auch diejenigen anderer endokriner Organe eine euploide Polyploidisierung zeigen (Biesterfeld et al. 1994), die hier bis zum 128-fachen des haploiden DNA-Gehaltes, also bis zu 128c, beschrieben (Heide 1982); dieses entspricht 2^6 , somit veranschaulichend netto dem Durchlauf von sechs Zellzyklen ohne Zellteilung. Außerdem lassen klinischen Arbeiten vermuten, dass DNA-Aneuploidie bei der Schilddrüse als Marker für Neoplasie verwendet werden könnte, nicht hingegen als Marker für Malignität. Dies gilt insbesondere für eine aus der hiesigen Abteilung stammende DNA-bildzytometrische Arbeit (Nadjari et al. 1999), in gewisser Weise aber auch für auswärtige DNA-durchflusszytometrische Studien (Grant et al., 1990, Kashyap et al., 1992, Francia et al. 1997, Horii et al. 1999).

1.5 Ziel der Arbeit

DNA-Messungen an Zellkernen zu diagnostischen und prognostischen Zwecken sind seit den 1960er-Jahren bekannt (Sandritter, 1964); erste methodische Überlegungen, z.B. durch die Gruppe um Caspersson, reichen noch länger zurück. Anfangs kamen dazu Mikroskop-Photometer zum Einsatz, die in den 1980er-Jahren durch immer wieder modernisierte und weiter entwickelte TV-Bildanalysesysteme ersetzt wurden.

State of the art ist, dass bei der DNA-Bildzytometrie die Auswahl zu messender Zellkerne durch den Untersucher erforderlich ist, indem dieser die Zellkerne mit einer Maus anklickt und somit den Messprozess und auch die Zuordnung der Zelle zu Referenz- oder Analysezellen festlegt. Hierdurch entsteht ein gewisser Zeitaufwand, der in der Routinediagnostik der Wirtschaftlichkeit der Methode entgegensteht (Böcking et al., 2015), unabhängig davon, ob ein Facharzt oder ein speziell ausgebildeter nicht-ärztlicher Mitarbeiter (MTA, CTA) die Untersuchung durchführt. Erst die Einführung von Zellkern-Klassifikatoren ("*Classifier*"), die von Zytopathologen gewebe- bzw. fixationsspezifisch trainiert werden (Friedrich, 2016), erlaubt es, nicht nur gesunde Referenzzellen automatisch zu detektieren, sondern auch Kerne malignitätsverdächtiger oder eindeutig maligner Zellen anhand ihrer spezifischen Morphologie (z.B. Dyskaryosen) mit einer Genauigkeit von ca. 93,5% zu erkennen (Friedrich et al., 2012).

Für die Entwicklung der Klassifikatoren und zur diagnostischen Erprobung wurde ein automatisch arbeitendes Bildanalysesystem für die DNA-Bildzytometrie entwickelt, dass im Funktionsbereich Cytopathologie zur Erprobung zur Verfügung steht. Es wurde vom chinesischen Mikroskophersteller Motic (Motic, Xiamen, China) in Zusammenarbeit mit dem Lehrstuhl für Bildverarbeitung der Rheinisch-Westfälischen Technischen Hochschule (RWTH) Aachen (Aach, Meyer-Ebrecht) und dem Institut / Funktionsbereich für Cytopathologie der Heinrich-Heine-Universität Düsseldorf (Böcking) entwickelt und in der hier vorliegenden Studie eingesetzt.

Ziel dieser Arbeit ist der Vergleich von Messungen mit dem neu entwickelten automatischen DNA-Bildzytometer mit manuellen DNA-bildzytometrischen Mes-

sungen aus der diagnostischen Routine an luftgetrockneten Ausstrichpräparaten von Feinnadelpunktaten der Schilddrüse. Die Reproduzierbarkeit und die Treffsicherheit des automatischen DNA-Bild-Zytometers mit dem Zellkern-Klassifikator für luftgetrocknete Präparate von Körperhöhlenergüssen wurde anhand von 40 zytologischen Präparaten von Schilddrüsenpunktionen untersucht.

Von allen untersuchten Präparaten liegt auch ein operatives histologisches Follow-up vor, so dass zusätzlich die Gesamttreffsicherheit, Sensitivität, Spezifität, positiver prädiktiver Wert und negativer prädiktiver Wert der konventionellen Zytologie sowie der manuellen und automatischen DNA-Bildzytometrie ermittelt werden soll.

2 Material und Methoden

2.1 Patienten und Präparate

Im elektronischen Befundarchiv des Funktionsbereichs Cytopathologie Düsseldorf dc-pathos (Fa. Dc-Systeme, Heiligenhaus) wurde in einem 10-Jahres-Zeitraum von 2004 bis 2014 nach Befunden von Ausstrichpräparaten nach Feinnadelpunktionen der Schilddrüse gesucht, bei denen im Rahmen der Routinediagnostik eine manuelle DNA-bildzytometrische Messung durchgeführt worden war. Da der Abgleich mit dem Operationspräparat angestrebt wurde, wurden aus der Vielzahl der unterschiedlichen Einsender aus Düsseldorf und durchaus auch der weiteren Umgebung in Nordrhein-Westfalen aus Gründen der Durchführbarkeit einige wenige Einsender ausgewählt. Die einbezogenen Präparate stammen hauptsächlich aus der Universitätsklinik Düsseldorf sowie aus der Gemeinschaftspraxis für Endokrinologie in Düsseldorf (Dr. Cissewski / Priv. Doz. Dr. Quadbeck / Prof. Dr. Tharandt).

Bei 92 Proben dieser Einsender wurde im oben genannten Zeitraum eine DNA-Bildzytometrie durchgeführt. Davon wurden 46 Schilddrüsenpräparate letztendlich in die vorliegende Studie einbezogen, da die Objektträger im Institut vorzufinden waren, noch eine gute Färbequalität aufwiesen und ein zeitnahes histologisches Follow-up erhoben werden konnte. Die histologischen Befunde wurden nachträglich im Archiv der Universitätsklinik Düsseldorf eingesehen bzw. in der Endokrinologischen Praxis erfragt bzw. eingesehen. Eine genauere Charakterisierung der 46 in die Studie aufgenommenen Fälle erfolgt in Kapitel 3.1.

2.2 Verarbeitung der Präparate für die Routinediagnostik

Die Präparate der Studie wurden nach Feinnadelpunktion vom einsendenden Kliniker ausgestrichen und luftgetrocknet (ca. eine Stunde bei Zimmertemperatur). Nach Versand und Eingang in der Cytopathologie wurden die Präparate anschließend routinemäßig nach May-Grünwald-Giemsa gemäß den dort geltenden *Standard Operation Procedures* (SOP) gefärbt. Dafür wurden sie zunächst 5 Minuten in einer Eosin-Methylenblau-Lösung (May-Grünwald-Lösung) inkubiert. Darauf folgten zwei Spülvorgänge in destilliertem Wasser, das über Zugabe einer Puffertablette auf pH = 6,8 eingestellt worden war. Eine Puffertablette (1,45 g), die für 1 l Aqua dest. ausgelegt ist, enthält 525,3mg Na₂HPO₄ und 857,4mg KH₂PO₄. Es folgten eine Inkubation in einer aus einer Stammlösung 1:20 mit Aqua dest. verdünnten, Azur-Eosin-Methylenblau-Lösung (Giemsa-Lösung) sowie zwei weitere Spülungen wie oben beschrieben. Die Färbelösungen wurden von der Firma Merck (Darmstadt) bezogen, ebenso die Puffertabletten. Schließlich wurden die Präparate unter Verwendung eines Eindeckmediums und eines Deckglases eingedeckt.

2.3 Zytologische Diagnostik

Die zytologische Diagnostik der MGG-gefärbten Präparate wurde in allen Fällen durch einen Facharzt für Pathologie vorgenommen, und zwar unter Anwendung der in der Standardliteratur niedergelegten zytologischen Kriterien. (Koss & Melamed, 2006) (Droese 1995). Dabei wurden die zytologischen Befunde nach der von den Deutschen Gesellschaften für Pathologie und Zytologie empfohlenen Befundklassifikation kategorisiert (siehe auch Kapitel 1.3.5), die aber für die Schilddrüsendiagnostik angepasst wurde. Hier kann naturgemäß „positiv“ nicht gleichbedeutend mit dem Nachweis eines malignen Tumors sein, sondern bedeutet in diesem Rahmen die Diagnose einer Neoplasie:

- **Neoplastische Zellen nachweisbar (= positiv):** Sämtliche Karzinome sowie die follikuläre Neoplasie.
- **Neoplastische Zellen wahrscheinlich (= dringender Verdacht):** Die diagnostischen Kriterien für eine sichere Diagnose einer Neoplasie sind entweder nicht hinreichend ausgeprägt oder es liegen zu wenige Tumorzellen vor, um eine sichere Diagnose zu stellen.
- **Neoplastische Zellen nicht sicher auszuschließen (= zweifelhaft):** Wenige atypische Zellen (z.B. abnorm aktivierte oder proliferierende Thyreozyten oder wenige stark degenerativ veränderte Tumorzellen). Außerdem werden hier Fälle eingeordnet, bei denen zahlreiche mikrofollikuläre Verbände nicht zwischen einer follikulären Neoplasie und einer zellreichen Kolloidstruma unterscheiden lassen.
- **Neoplastische Zellen nicht nachweisbar (= negativ):** In diese Kategorie fallen Kolloidstrumen, hyperplastische Knoten und Entzündungen mit oder ohne regressive Veränderungen, d.h. Fälle, bei denen neoplastische Zellen nicht nachweisbar sind. Es liegen ausschließlich normale und/oder proliferierende und/oder reaktiv veränderte Zellen vor.
- **Unzureichendes Material:** Präparate mit nicht repräsentativer Zellzahl oder ohne Follikelepithelien (z.B. lediglich Zellen aus dem Punktionsweg, wie quergestreifte Muskulatur).

2.4 Feulgen-Färbung

Nach konventioneller zytologischer Routinediagnostik und Befunderstellung wurden sämtliche Präparate durch Fachärzte des Funktionsbereiches Cytopathologie einer zusätzlichen diagnostischen Begutachtung unterzogen. Alle zytologisch diagnostisch relevanten Regionen wurden auf dem Deckglas mit einem Farbstift markiert. Da das

Deckglas im Rahmen der Feulgen-Färbung vom Präparat abgelöst werden musste, wurde von den Präparaten jeweils eine Fotokopie ihrer Oberfläche angefertigt, so dass die Kontur der Farbmarkierungen dokumentiert war. Im Anschluss an die Feulgen-Färbung wurden diese Areale auf dem neuen Deckglas nachgezeichnet. Dadurch wurde sichergestellt, dass mit der DNA-Bildzytometrie die gleichen Zellgruppen gemessen wurden, die bereits für die zytologische Diagnose von Relevanz waren.

Grundlage für die Messung des DNA-Gehaltes von Zellkernen ist die quantitative Färbung der Zellkern-DNA nach Feulgen und Rossenbeck (Feulgen & Rossenbeck, 1924), modifiziert nach Graumann (Graumann, 1953). Die ursprünglich nach May-Grünwald-Giemsa vorgefärbten Objektträger müssen zunächst mindestens 72 Std. in Xylol entdeckelt und anschließend in einem Methanolbad gewaschen werden. Im Anschluss folgt die Feulgen-Färbung (siehe Tab.1) in einem mit einer heizbaren Küvette ausgestatteten und in einigen anderen Bauteilen speziell modifizierten Färbeautomaten (Shandon-Varistain® 24-4) (Chatelain et al., 1989).

Färbeschritt	Küvetteninhalt	Inkubationszeit
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	5N HCL (ca. 27°C)	60 min
8	Aqua dest.	2 min
9	Aqua dest.	2 min
10	Aqua dest.	2 min
11	Schiff'sches 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	99% Ethanol	10min
19	99,5% Ethanol	10 min
20	Xylol	15 min

Tabelle 1: Ablauf der Feulgen-Färbung

Die histochemische Reaktion läuft in zwei Reaktionsschritten ab. Das Prinzip der Feulgen-Färbung beruht darauf, dass durch saure Hydrolyse mit 5N HCl eine Zerstörung der glykosidischen Bindung der Purinbasen an der C1-Stelle erfolgt, so dass freie Aldehydgruppen an den Desoxyriboseanteilen der DNA entstehen. Anschließend reagieren diese mit der aromatischen Aminogruppe des Schiff'schen Reagenz (Schiff'sches Reagenz mit Pararosanilin (rot) oder Thionin (blau)) zu einem Farbkomplex, welcher eine Eigenfärbung der DNA bewirkt. Die Zellkerne erhalten so eine violette Färbung, das Cytoplasma bleibt aufgrund fehlender DNA ungefärbt. RNA wird ebenfalls nicht angefärbt, weil die Spaltung der Glykosidbindung an der C1-Stelle nicht möglich ist.

Nach Abschluss der Feulgen-Färbung erfolgte die manuelle Eindeckelung der Präparate mit einem konventionellen Medium für die Zytologie und Histologie, z. B. VitroClud, und einem üblichen Deckglas. Vom DNA-bildzytometrischen Einsendeschein wurden dann die oben erwähnten aufkopierten Markierungen manuell auf das neue Deckglas übertragen. Das Präparat wurden bis zur Messung dunkel bei Zimmertemperatur aufbewahrt, um ein Ausbleichen zu verhindern.



Abbildung 1: Fertige, nach Feulgen gefärbte Präparate

2.5 Geräte und Messtechnik

Alle für diese Studie verwendeten Schilddrüsenzytologien wurden in der primären Diagnostik (Zeitraum 2004-2014) zunächst mit dem jeweils im Funktionsbereich Cytopathologie in der Routinediagnostik eingesetzten interaktiven DNA-bildzytometrischen Messsystem, zuletzt der *MotiCyte DNA Manual Cytology Workstation* (Motic, Xiamen, China), gemessen. Die Messung mit dem neu entwickelten

automatischen Messsystem *MotiCyte DNA Cytology Workstation auto* (Motic, Xiamen, China) wurden an denselben Präparaten nachträglich im Jahr 2015 durchgeführt.

2.6 Prinzip der DNA-Bildzytometrie

Das Prinzip bei der DNA-Bildzytometrie ist die densitometrische Bestimmung der integrierten optischen Dichte (IOD) von Zellkernen. Die IOD ist die Summe der Grauwerte aller Pixel eines Zellkernbildes.

Als Referenzwert für einen normalen diploiden Chromosomensatz wird die mittlere IOD von regelrechten Zellen, sogenannten „Referenz-“ oder „Eichzellen“, verwendet. Sie haben definitionsgemäß einen DNA-Gehalt von $2c$ (c = content = DNA-Gehalt). Dies ist das DNA-bildzytometrische Äquivalent eines diploiden Chromosomensatzes.

Von der Bildanalysesoftware wird für jede Messung eine vom Untersucher korrigierbare Messmaske generiert, welche das Bild eines Zellkerns exakt abdeckt. Darin wird die IOD bestimmt. Die IOD der Analysezellen wird mit der IOD einer Population von morphologisch unauffälligen Referenzzellen aus dem gleichen Präparat verglichen. Durch den Vergleich mit dem diploiden Standard von $2c$ kann somit jeder Analysezelle ein konkreter c -Wert zugeordnet werden.

Eine zweite Messmaske dient der Bestimmung des mittleren Grauwertes des Hintergrundes, auch Leerwert (I_0) genannt. Die Bildanalysesoftware korrigiert insbesondere Streulicht (glare) und Lichtbeugungen am Zellkernrand, um Messfehlern vorzubeugen (Kindermann & Hilgers, 1994).

2.7 Interaktive (manuelle) DNA-Bildzytometrie

Für die interaktive Messung wurde in der primären Diagnostik die *MotiCyte DNA Manual Cytology Workstation* eingesetzt. Dieses TV-Bildanalysesystem erfüllt die Richtlinien der *Consensus Reports* der European Society for Analytical Cellular Pathology (ESACP) über die Standardisierung der diagnostischen DNA-Bildzytometrie (Böcking et al., 1995), (Giroud et al., 1998), (Haroske et al., 1998) (Haroske et al., 2001).

Das PC-gestützte Bildanalysesystem besteht aus (Abb. 2)

- einem Motic BA400 Mikroskop (Motic, Xiamen, China) mit verschiedenen Objektiven, darunter ein 40er Objektiv für den eigentlichen Messvorgang,
- einer am Mikroskop angeschlossenen 12-bit CCD-Kamera mit einer Auflösung von 1360x1024 Pixel (MotiCyte285A; Motic Xiamen, China),
- einem PC mit MotiCyte-DNA cytometry Software 2.3. und dem zugehörigen Klassifikator sowie
- einem Farbdrucker.

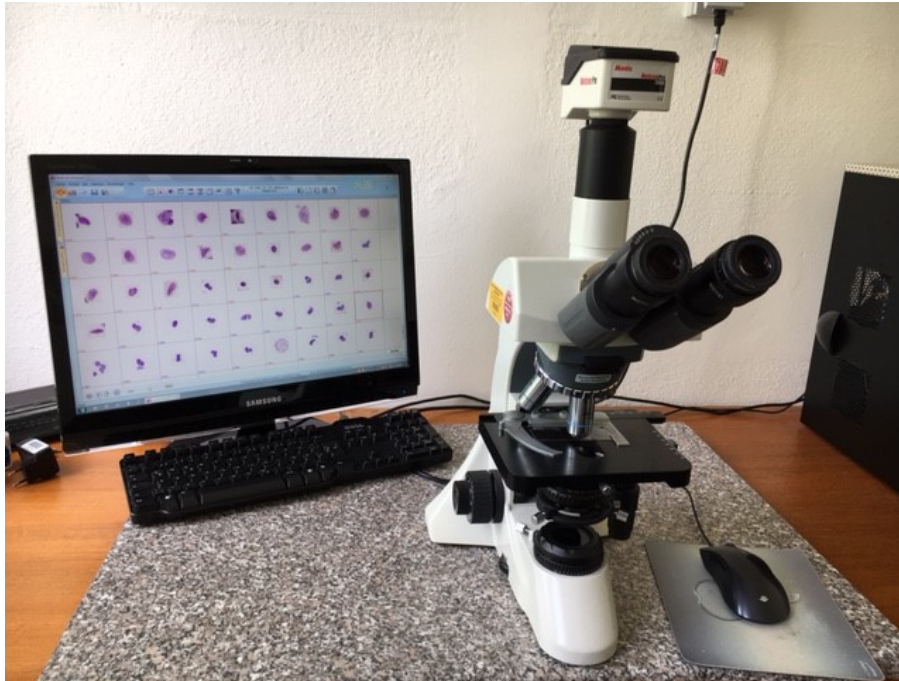


Abbildung 2: Interaktives DNA-Bildzytometer MotiCyte DNA

2.7.1 Messvorgang

Sämtliche Messergebnisse der interaktiven Zytometrie, die in der vorliegenden Studie verwendet wurden, stammen aus der diagnostischen Routine des Funktionsbereichs Cytopathologie und wurden von eigens geschulten Medizinisch-Technischen Assistentinnen bzw. Fachärzten für Pathologie durchgeführt.

Vor jeder neuen Messung eines Präparates ist darauf zu achten, das Mikroskop neu zu kalibrieren.

Die Messung (Abb. 3) erfolgt interaktiv durch den Untersucher. Das Präparat wird zunächst auf den Objektisch gelegt und das Mikroskop sowie die Software auf das zu messende Präparat kalibriert. Dazu wird ein zellreicher Ausschnitt aus dem Präparat genau fokussiert. Der Vorgang kann parallel auf dem Computer verfolgt werden. Nach Einstellung von Blende und Kondensor sowie Justierung des Mikroskops nach dem Köhler'schen Beleuchtungsprinzip zur Streulichtreduktion erfolgt eine Leerwertmessung. Dafür wird ein leeres Feld aus dem Präparat aufgesucht und sein mittlerer Grauwert berechnet (Weißabgleich). Anschließend ermöglicht eine Messung bei geschlossenem Strahlenteiler am Mikroskop einen Schwarzabgleich, da kein Bild mehr zur Kamera gelangt.



Abbildung 3: Messmonitor des interaktiven Messsystems, hier mit einem Mundschleimhautabstrich

Die zu messenden Zellkerne, welche auf dem Monitor mit der Maus einzeln angeklickt werden können, werden vom Untersucher nach morphologischen Kriterien ausgewählt. Die Software generiert über der angeklickten Zelle die Messmaske, in der die IOD gemessen wird. Eine Bildergalerie (*image gallery*) führt die Zellkerne der Analysezellen und der Referenzzellen getrennt auf. Unscharfe, überlagerte oder autolytische Zellkerne sowie Artefakte können durch den Untersucher aus der Bildergalerie gelöscht werden. Bei der interaktiven Messung wählt der Untersucher, wenn möglich, ca. 300 Analysezellen aus. Als Referenzzellen dienen morphologisch unauffällige, normale, im Untersuchungsmaterial des Patienten vorhandene Zellen, wie Lymphozyten, Granulozyten, oder normale Epithelzellen, von denen mindestens 30 gemessen werden sollten. Bei der Messung der Referenzzellen ist zur Sicherstellung des Qualitätsstandards darauf zu achten, dass der Variationskoeffizient der IOD's (CV) unter 5% liegt, damit die Messungen zur Kalibrierung eines DNA-Histogramms verwendet werden können (Böcking, 1995) (Giroud et al., 1998). Außerdem muss der Korrelationskoeffizient zwischen IOD und der Fläche der Referenzzellen gemäß den Anforderungen der ESACP unterhalb von 0,4 liegen, d.h. dass keine Abhängigkeit der IOD normaler Referenzzellen von der Fläche der Zellkerne bestehen darf (Haroske et al., 2001).

2.7.2 Datenauswertung und Dokumentation

Die Ergebnisse der Messung werden in einem dreiseitigen Ausdruck dokumentiert (Abb. 4 - 6.), der sowohl methodische Aspekte, als auch das klinische Messergebnis zusammenfasst. Die Seite 1 des Ausdrucks wird dem klinischen Einsender des Falles als Diagnoseblatt übersandt.

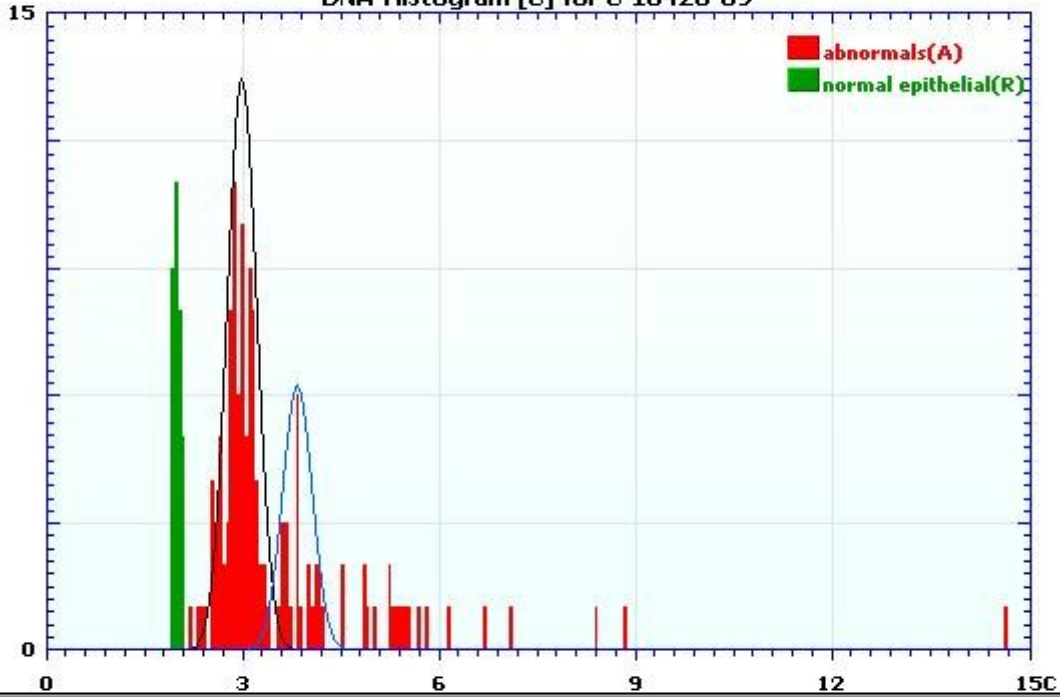
Cytopathologie. Universitätsklinikum Düsseldorf Leitung: Prof. Dr. med. S. Biesterfeld Moorenstr. 5, 40225 Düsseldorf, Tel.: (0211)-8118346, Fax: (0211)-8118402			
DNA Image Cytometry			
Patient ID:	C 16426-09	Material ID:	C 16426-09
Last Name:			
First Name:			
Birth Date:	1-Jan-1970	Date of sample:	28-Sep-2011
Sex:	Male	Staining:	Feulgen
Barcode: 			
DNA-Histogram [C] for C 16426-09			
			
Diagnostic Interpretation of DNA-Distribution			
DNA-Ploidy:			
DNA-Grading:			
Recommendation:			
Comment:			
Physician's Signature		Date	
Technician's Signature:		Date	
MotiCyte-DNA-Manual Report			

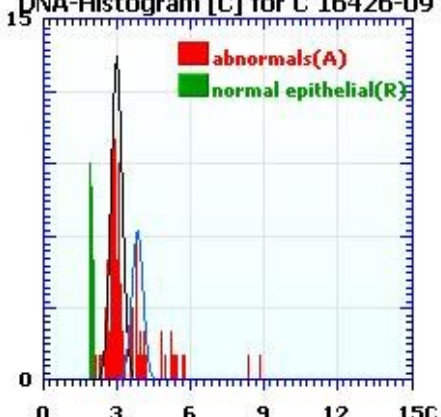
Abbildung 4: Seite 1 des dreiseitigen Berichts einer manuellen DNA-zytometrischen Messung

Cytopathologie. Universitätsklinikum Düsseldorf
Leitung: Prof. Dr. med. S. Biesterfeld
Moorenstr. 5, 40225 Düsseldorf, Tel.: (0211)-8118346, Fax: (0211)-8118402

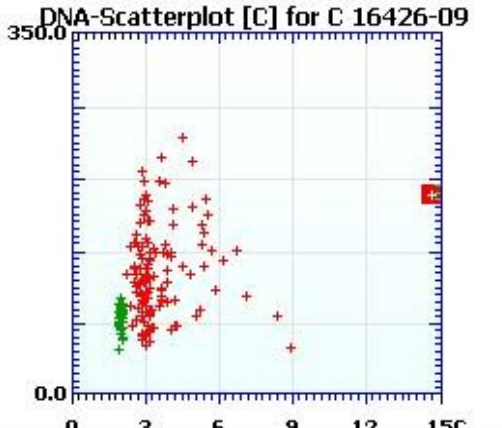
DNA Image Cytometry

Patient ID: C 16426-09 Material ID: C 16426-09
Last Name: _____
First Name: _____
Birth Date: 1-Jan-1970 Date of sample: 28-Sep-2011
Sex: Male Staining: Feulgen
Barcode: 

DNA-Histogram [C] for C 16426-09



DNA-Scatterplot [C] for C 16426-09



Statistics	Stem1 1	Stem1 2
Modal Value	2.97	3.82
DNA Index	1.49	1.91
Correlation	-0.18	-0.20

Reference Cell Type	Basic Indices	
normal epithelial	2c Reference IOD	0.00
Number	2c Deviation Index	3.85
Correction	Diploid Dev. Quotient	1.80
CV[%]	Z Value	20.12
	Ploidy Balance	-0.32

Analysis Cells		
Number	131	Aneuploid R.(d) 0.00
Min.[c]	2.23	Aneuploid R.(t) 0.00
Max.[c]	14.66	Modal[c] 2.90

Absolute Indices(events)		Relative Indices (rate)	
2.5c Ex Events	127	2.5c Ex. Rate	96.95
3c Ex. Events	81	3c Ex. Rate	61.83
4c Ex. Events	29	4c Ex. Rate	22.14
5c Ex. Events	16	5c Ex. Rate	12.21
7c Ex. Events	4	7c Ex. Rate	3.05
9c Ex. Events	1	9c Ex. Rate	0.76

Diagnostic Interpretation of DNA-Distribution

Physician's Signature

Date

Technician's Signature:

Date

MotiCyte-DNA-Manual Report

Abbildung 5: Seite 2 des dreiseitigen Berichts einer manuellen DNA-zytometrischen Messung

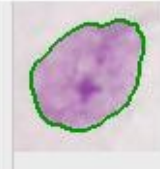
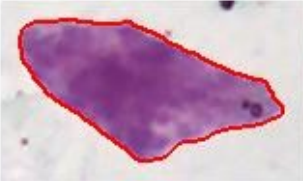
Cytopathologie, Universitätsklinikum Düsseldorf			
Leitung: Prof. Dr. med. S. Biesterfeld			
Moorenstr. 5, 40225 Düsseldorf, Tel.: (0211)-8118346, Fax: (0211)-8118402			
DNA Image Cytometry			
Patient ID:	C 16426-09	Material ID:	C 16426-09
Last Name:			
First Name:			
Birth Date:	1-Jan-1970	Date of sample:	28-Sep-2011
Sex:	Male	Staining:	Feulgen
Barcode:			
Reference cells - normal epithelial			
			
2.11c	2.11c	2.11c	2.09c
Analysis cells			
			
14.66c	8.88c	8.38c	
			
7.10c	6.69c	6.18c	
MotiCyte-DNA-Manual Report			

Abbildung 6: Seite 3 des dreiseitigen Berichts einer manuellen DNA-zytometrischen Messung

2.8 Automatische DNA-Bildzytometrie

Bei dem automatischen Bildzytometer (Abb. 7) handelt es sich um das Motic BA400 microscope™ (Motic, Xiamen, China). Es ist mit einer MotiCamPro 285A Kamera™ (Motic, Xiamen, China) mit einer Auflösung von 1360x1024 Pixeln und der MotiCyt DNA image cytometry software™ (Motic, Xiamen, China) ausgestattet.

Die Mess-Software wurde in einer Kooperation zwischen der Firma Motic, dem Institut für Cytopathologie der Heinrich Heine-Universität Düsseldorf und dem Lehrstuhl für Bildverarbeitung der RWTH Aachen entwickelt (zusammengefasst bei Friedrich 2016). Sie basiert auf Gewebe- und Fixations-adaptierte Zellklassifikatoren, die so konzipiert wurden, dass sie in der Lage sind, verschiedene Zelltypen zu erkennen und ihre Messwerte getrennt zu erfassen, so dass als Analysezellen nur mikroskopisch verdächtige Zellen herangezogen werden.

Für die hier vorliegende Arbeit an Ausstrichpräparaten von Schilddrüsenpunktaten wurde ein eigentlich für luftgetrocknete Präparate von Körperhöhlenergüssen entwickelter Zellklassifikator verwendet, da für die Schilddrüse noch kein eigener Klassifikator existiert. Er wurde gewählt, weil auch Schilddrüsenpunktate luftgetrocknet werden, was Schrumpfungen der Zellkerne vorbeugt, und weil die Zellkerne der Thyreozyten in Größe und Kontur denen von Mesothelien stark ähneln.

Die übrigen bisher entwickelten Klassifikatoren betreffen Prostatakarzinome nach Zellvereinzelung sowie Urinzentrifugate nach Alkohol-Fixation sowie Alkohol-fixierte plattenepitheliale Ausstriche aus der Gynäko- und Oralpathologie und waren für diese Studie ungeeignet.

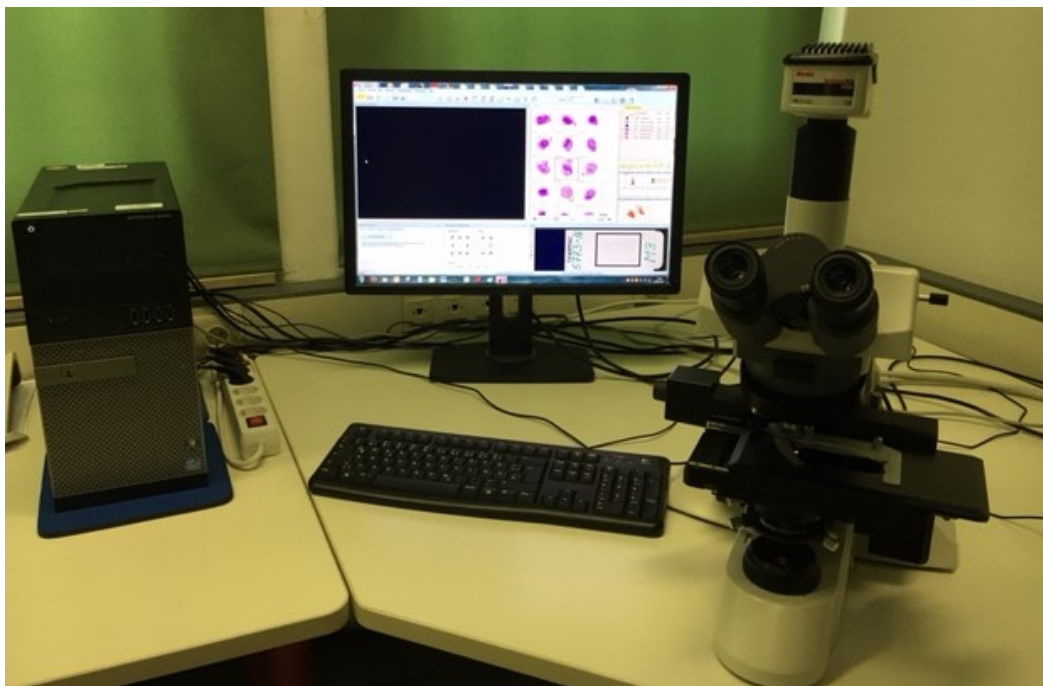


Abbildung 7: Automatisches DNA-Bildzytometer (Moticyte DNA Cytology Workstation auto)

2.8.1 Scannen der Präparate

Eingangs wird eine Übersichtskarte des zu scannenden Präparates erzeugt. Dafür wird der Objektträger auf dem Objektisch des motorisierten Mikroskops aufgelegt und bei niedriger Primärvergrößerung automatisch gescannt. Es erscheint eine Übersichtskarte des gesamten Präparats auf dem Bildschirm (Abb. 8).

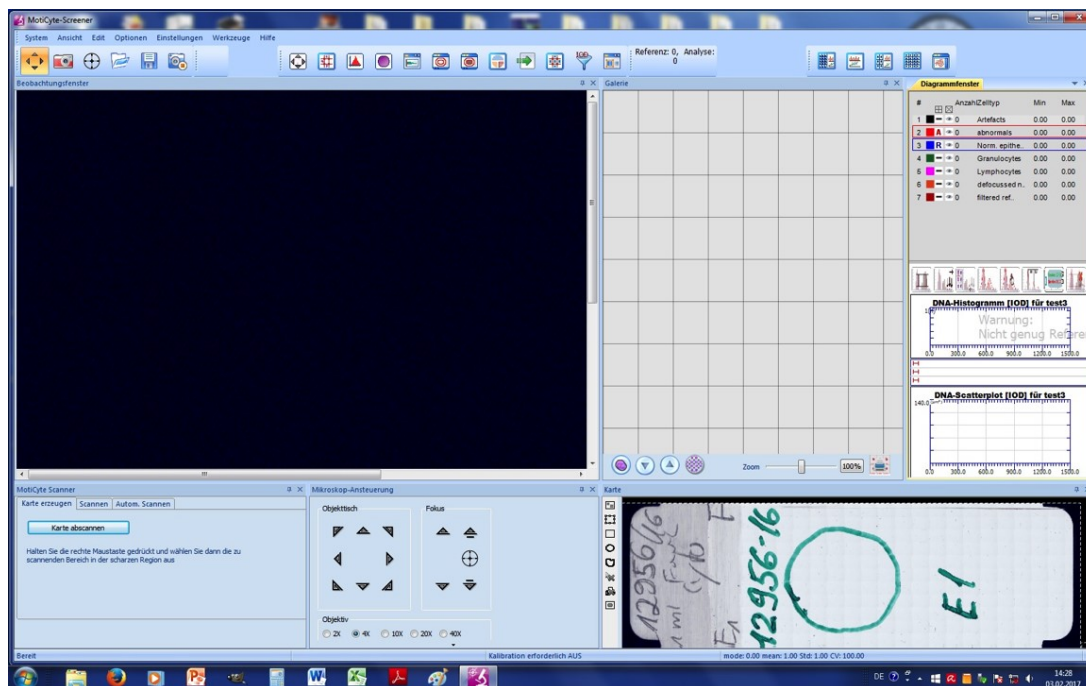


Abbildung 8: Monitordarstellung der erzeugten Übersichtskarte des Objektträgers (rechts unten)

Auf dem so entstandenen Bild können dann die diagnostisch relevanten Bereiche mit mikroskopisch auffälligen Zellen digital markiert werden, in welchen ein systematisches Scannen bei hoher Vergrößerung (40x-Objektiv) erfolgen soll. Das manuelle Hinzufügen von Fokuspunkten dient dem Zweck, dass auch in weiter entfernten Arealen des Präparates die zu messenden Zellen nicht außerhalb des Fokusumfanges des Autofokus geraten und ist Grundvoraussetzung für ein automatisches systematisches Scannen innerhalb dieses Areals. Außerdem können wichtige Areale eingegrenzt und einzeln gemessen werden (Abb. 9).

Die gewünschte Vergrößerung des Präparats wird genau wie die Höhe des Objektisches automatisch eingestellt. Es folgt die manuelle Kalibrierung des Zytometers, wie bereits in Kapitel 2.5 beschrieben. Durch Anklicken des Feldes „Karte abschnitten“ wird das Präparat in dem markierten Bereich gescannt. Der Messvorgang findet dann automatisch statt und kann auf dem Monitor mitverfolgt und, wenn nötig, unterbrochen werden.

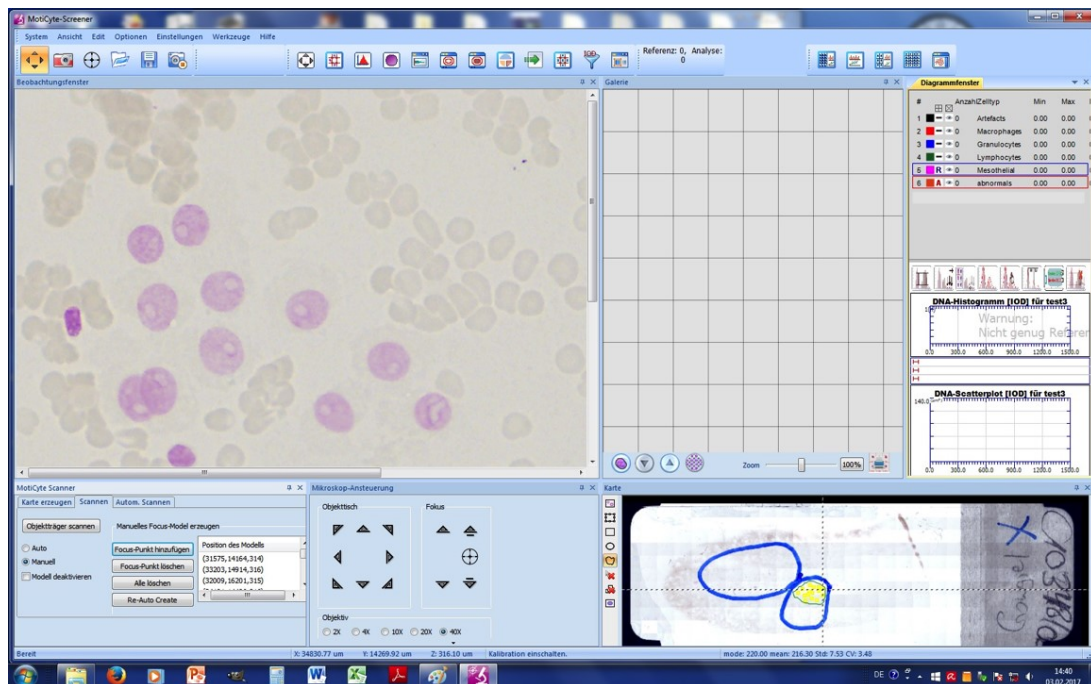


Abbildung 9: Monitordarstellung nach Setzen der Fokuspunkte (gelb) im zellhaltigen Messareal (grün)

2.8.2 Digitale Klassifizierung

Zunächst werden sämtliche gefärbten Objekte in ihrer exakten Form und Größe identifiziert und mit hochspezifischen Masken versehen, innerhalb derer die weitere Bildanalyse und gegebenenfalls DNA-Messung erfolgt (Böcking et al., 2015).

Für die Messungen an luftgetrockneten Schilddrüsenpunktaten wurde mangels eines eigenen Klassifikators für Schilddrüsenpunktate der eigentlich für Körperhöhlenergüsse entwickelte Zellklassifikator verwendet. Da die Kriterien für Kernatypien aber bei beiden luftgetrockneten Präparationen ähnlich sind, wurde dieses Vorgehen gewählt. Der Erguss-Klassifikator kategorisiert neben Lymphozyten, Granulozyten, Makrophagen, Artefakten, defokussierten Zellen und Analysezellen allerdings noch normale Mesothelien, die in der Schilddrüse nicht vorkommen. Die als Mesothelien klassifizierten Zellen wurden nach der Messung zu den Analysezellen hinzugefügt.

In Abbildung 10 ist der Übersichtsbildschirm nach Ende eines Scanvorgangs gezeigt. Links sieht man das Beobachtungsfenster, links unten kann der Scanvorgang und das Mikroskop gesteuert werden. In der Mitte erscheint die Bildergalerie mit Darstellung von 12 der erfassten Analysezellen. Rechts oben ist das Diagrammfenster zu sehen, in dem die einzelnen Kategorien, das DNA-Histogramm und das DNA-Scatterplot dargestellt sind. Jede der einzelnen Teilansichten kann auch gesondert dargestellt werden.

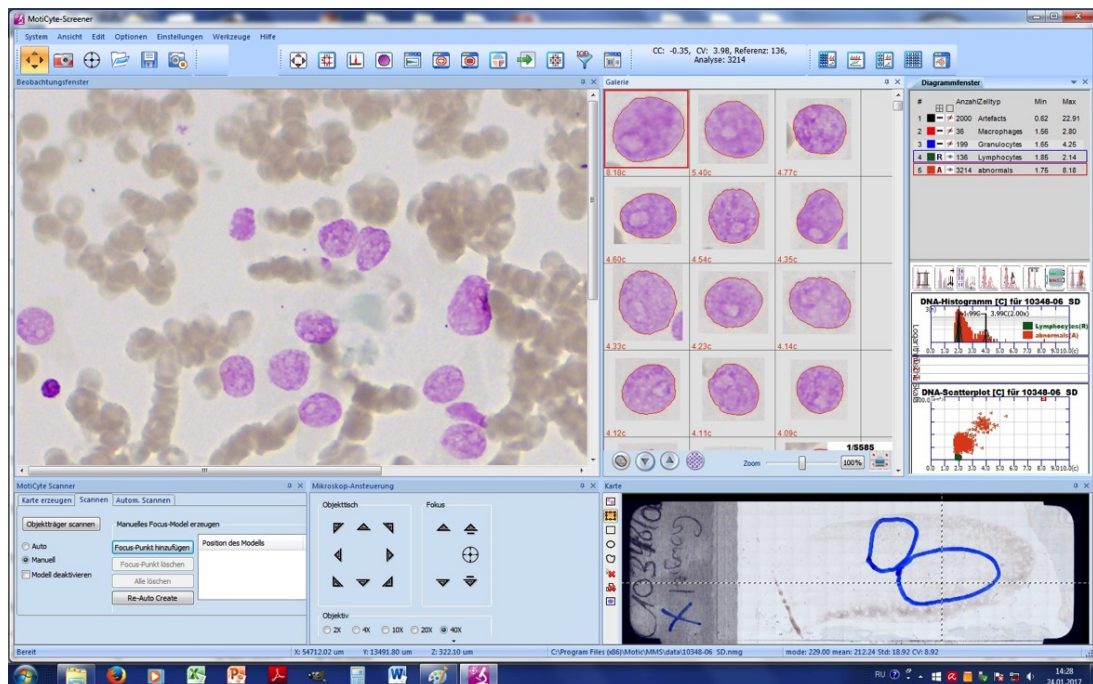


Abbildung 10: Monitordarstellung des fertigen Scanvorgangs

2.8.3 Datenauswertung und Dokumentation

Die Ergebnisse der Messungen wurden pro Präparat in drei Versionen gespeichert:

1. **Rohdaten:** unmittelbar nach automatischer Messung, ohne Bearbeitung oder manueller Umklassifizierung der einzelnen Zellen (Abb. 11).

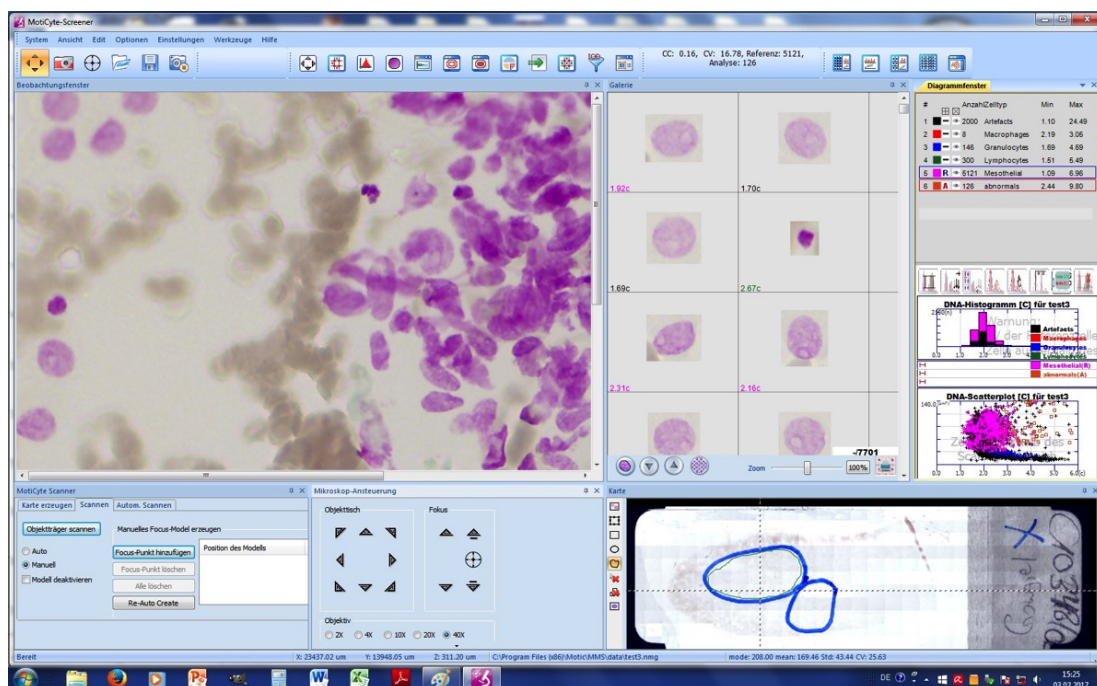


Abbildung 11: Rohdaten eines gescannten Präparates

2. nach **manueller Umklassifizierung** (Abb. 12) der Gruppe „Mesothelien“ zu den Analysezellen und Kontrolle der **Eichung**, d.h. nach Auswahl der Referenzzellen und automatischer Filterung zum Erreichen der zulässigen Werte für CV und CC.

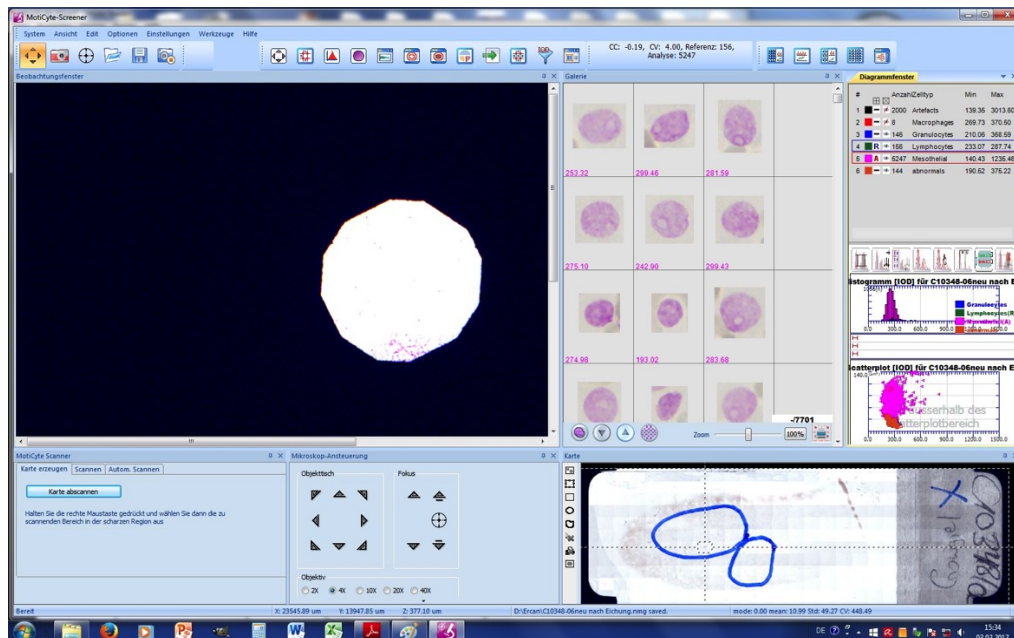


Abbildung 12: Präparat nach Umklassifizierung

3. nach **Bearbeitung der Bildgalerie**: Löschen von Dubletten, Überprüfung der *image gallery* der einzelnen Kategorien auf Richtigkeit, Durchsuchung der Artefakte oder defokussierten Zellen nach eventuell automatisch nicht detektierten abnormalen Zellen oder Thyreozyten. Erst durch Nachbearbeitung der Bildgalerie können die automatischen Messungen mit der manuellen Messung verglichen werden (Abb. 13).

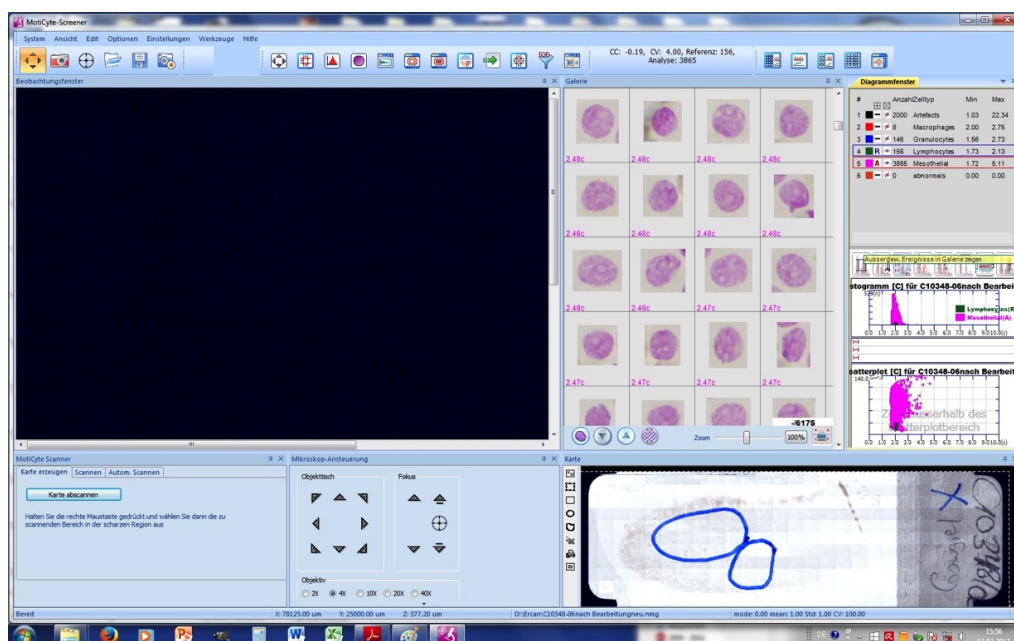


Abbildung 13: Präparat nach Bearbeitung

Nach dem Messvorgang werden alle gescannten Zellen in einer Bildergalerie aufgeführt (Abb. 14), in der nachträglich Zellen bearbeitet werden können. Dort können gegebenenfalls die Segmentierung (d.h. automatische Trennung eng zusammenliegender Zellkerne) der Objekte manuell korrigiert (z.B. Trennen von Dubletten), diese manuell umklassifiziert oder gelöscht werden.

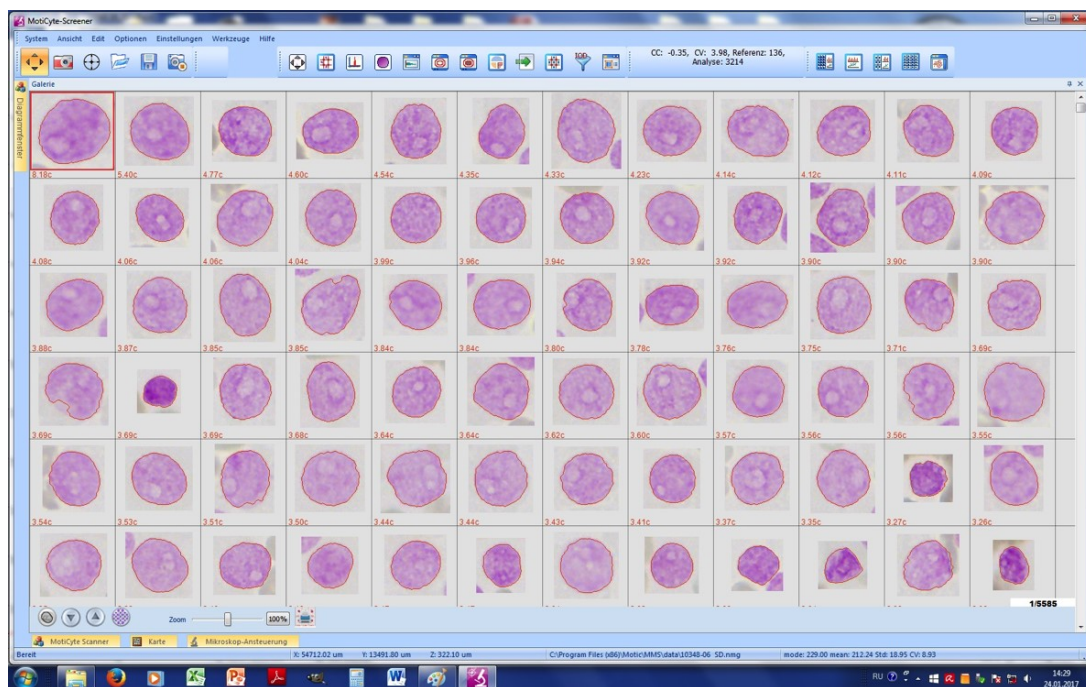


Abbildung 14: Bildergalerie der Analysezellen nach Abschluss aller Bearbeitungen

Das Ergebnis der jetzt vom Untersucher gesichteten und ggf. bearbeiteten Messung wird anschließend im DNA-Histogrammen und Scatter-Plot dargestellt. Am Ende jedes Scans erscheinen ggf. Warnmeldungen gemäß den Qualitätsstandards der ESACP, beispielsweise bei zu wenig gemessenen Referenz- oder Analysezellen sowie zu hohen oder zu niedrigen CC- oder CV-Werten der Referenzzellen. Diese Probleme können, z.B. durch automatisches Filtern der Referenzzellen zum Erreichen der empfohlenen Grenzwerte für CC oder CV, behoben werden.

Pro Messung wird ein dreiseitiger Bericht (Abb. 15-17, dort wiedergegeben nach Löschung aller klinischer Daten) ausgegeben, welcher neben dem DNA-Histogramm und einem Scatter-Plot repräsentative Abbildungen gemessener Referenz- und Analyse-Zellkerne sowie standardisierte Indizes der DNA-Verteilung und die von einem Pathologen autorisierte diagnostische bzw. prognostische Interpretation der Messergebnisse enthält.

Die in Abb. 15 dargestellte erste Seite stellt dabei den eigentlichen Befund dar, welcher dem Einsender übermittelt wird.

Cytopathologie - Universitätsklinikum Düsseldorf
 Leitung: Dr.(UA) N. Pomjanski, Dr. M. Schramm
 Moorenstr. 5, 40225 Düsseldorf; Tel.: (0211) - 81 18346, Fax: (0211) - 81 18402

DNA-Bildzytometrie

Patienten ID: _____ Material ID: _____
 Vorname: _____ Organ: _____
 Nachname: _____ Material: _____
 Geburtsdatum: _____ Datum der Probe: _____
 Geschlecht: _____ Färbung: _____
 Klinische Angaben: _____

DNA-Histogramm [C] für C10400-06

DNA-Ploidie: _____
 DNA-Klasse: _____
 Empfehlung: _____

Unterschrift Arzt/Ärztin: Ercan _____ 14/12/2015
 Unterschrift ZTA/MTA: _____ 14/12/2015

Abbildung 15: Seite 1 des dreiseitigen Berichts einer automatischen DNA-zytometrischen Messung

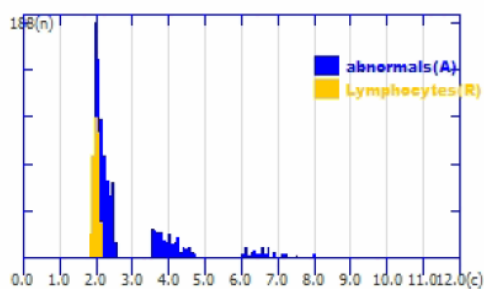
Cytopathologie - Universitätsklinikum Düsseldorf
 Leitung: Dr.(UA) N. Pomjanski, Dr. M. Schramm
 Moorenstr. 5, 40225 Düsseldorf; Tel.: (0211) - 81 18346, Fax: (0211) - 81 18402

DNA-Bildzytometrie

Patienten ID:	Material ID:
Vorname:	Organ:
Nachname:	Material:
Geburtsdatum:	Datum der Probe:
Geschlecht:	Färbung:

Statistik

DNA-Histogramm [C] für C10400-06



Referenzzellen		Lymphocytes	
Anzahl	344		
CC	-0.10	CV	3.88
2c Referenz-IOD			321.56
Analysezellen		abnormals	
Anzahl	1009	Aneuploid.E(d)	886
Min.[c]	1.96	Aneuploid.E(t)	163
Max.[c]	9.95	Modal-Wert [c]	1.97
2c Deviation Index			45.6
Mittlere Ploidie[C]			3.35
Absolute Indizes		Relative Indizes	
2.5c Ex. Events	295	2.5c Ex. Rate	29.24
3c Ex. Events	284	3c Ex. Rate	28.15
4c Ex. Events	163	4c Ex. Rate	16.15
5c Ex. Events	78	5c Ex. Rate	7.73
7c Ex. Events	24	7c Ex. Rate	2.38
9c Ex. Events	1	9c Ex. Rate	0.10

	Peak A	Peak B	Peak C
Modalwert	2.05c	4.18c	0.00c

Zellfraktionen: Statistik

Fraktion	Min	Max	Mittelwert	Modalwert	CV	Anzahl
1	—	—	—	—	—	—
2	—	—	—	—	—	—
3	—	—	—	—	—	—

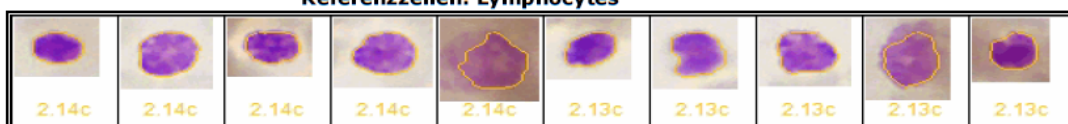
Abbildung 16: Seite 2 des dreiseitigen Berichts einer automatischen DNA-zytometrischen Messung

Cytopathologie - Universitätsklinikum Düsseldorf
 Leitung: Dr.(UA) N. Pomjanski, Dr. M. Schramm
 Moorenstr. 5, 40225 Düsseldorf; Tel.: (0211) - 81 18346, Fax: (0211) - 81 18402

DNA-Bildzytometrie

Patienten ID:	Material ID:
Vorname:	Organ:
Nachname:	Material:
Geburtsdatum:	Datum der Probe:
Geschlecht:	Färbung:

Referenzzellen: Lymphocytes



Analysezellen: abnormal

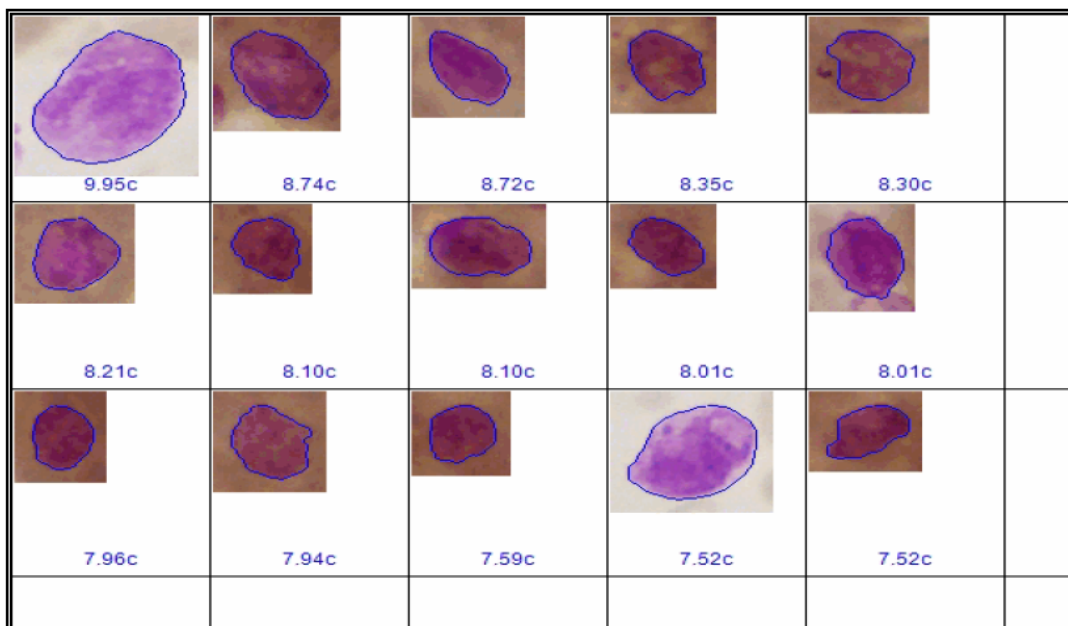


Abbildung 17: Seite 3 des dreiseitigen Berichts einer automatischen DNA-zytometrischen Messung

2.8.4 Vergleichsparameter

Von den in den DNA-zytometrischen Befundberichten (siehe Abb. 5 und Abb. 16) dokumentierten Größen wurden die im Folgenden genannten im Ergebnisteil berücksichtigt; die Definitionen folgen dabei den Vorgaben der ESACP-Konsensuskonferenz in Oslo aus dem Jahr 1997 (Haroske et al., 1998) bzw. einem Update hierzu aus dem Jahr 2001 (Haroske et al., 2001).

- **Zahl der Referenzzellen**
- **IOD (Integrierte optische Dichte) der Referenzzellen:** Mittelwert der aufsummierten Extinktionswerte der einzelnen Pixel der Referenzzellen; dieser Wert wird, die Erfüllung von Qualitätskriterien vorausgesetzt, als 2c-Äquivalent der Messzellen angenommen.
- **Zahl der Analysezellen**
- **Minimaler DNA-zytometrischer Messwert der Analysezellen**
- **Maximaler DNA-zytometrischer Messwert der Analysezellen**
- **5cEE:** Messwerte mit Werten oberhalb von 5c, d.h. oberhalb tetraploider Werte (4c) zuzüglich großzügiger Marge für Messfehler
- **5cER:** Prozentualer 5cEE-Wert, bezogen auf die Messzellzahl
- **9cEE:** Messwerte mit Werten oberhalb von 9c, d.h. oberhalb oktoploider Werte (8c) zuzüglich großzügiger Marge für Messfehler
- **9cER:** Prozentualer 9cEE-Wert, bezogen auf die Messzellzahl
- **DNA-Stammlinie:** Höchster Peak der Messzellpopulation im Histogramm. Der Begriff DNA-Stammlinie (Seidel & Sandritter, 1963) beschreibt die Lage der G₀/G₁-Fraktion einer Zellpopulation und sollte mit einer kleineren Zellpopulation mit einem doppelt so hohen DNA-Gehalt ("Verdopplungsgipfel") einhergehen, der der G₂/M-Fraktion entspricht. Laut ESACP-Protokoll wird eine DNA-Stammlinie im Histogramm als ein signifikantes lokales Maximum, das kein Duplikat eines anderen Peaks darstellen darf, definiert.

2.8.5 Diagnostische Interpretation

Zur diagnostischen Interpretation der Messungen wurden die DNA-Stammlinie und der 9cEE herangezogen und gemäß den ESACP-Kriterien (Haroske et al., 2001) als "aneuploid" bzw. "euploid" beurteilt:

- **DNA-Stammlinie ("Stammlinieninterpretation"):** Eine DNA-Stammlinie im Bereich von 2c +/- 10%, also bei 1,80c bis 2,20c, wird als euploid (diploid) beurteilt. Bei einem Wert bei 4c +/- 10%, entsprechend von 3,60c bis 4,40c, wird das Ergebnis als euploid (tetraploid) interpretiert. Stammlinien außerhalb dieser Bereiche gelten als aneuploid.

- **9cEE ("Einzelzellinterpretation"):** Bei Nachweis von Messwerten mit Werten oberhalb von 9c ($9cEE > 0$) gilt die Messung als aneuploid, bei Fehlen solcher Zellen ($9cEE = 0$) als euploid.

2.9 Statistische Untersuchungen

Die statistischen Untersuchungen wurden mit Microsoft Excel sowie der Software IBM SPSS Statistics 23 am Funktionsbereich Cytopathologie in Kooperation mit der Statistikerin Isabela Karoline de Santana Almeida Araújo erstellt und ausgewertet.

Für die deskriptive Statistik standen alle üblichen Daten wie Mediane, Mittelwerte, Varianz und Standardabweichung zur Verfügung. Ferner wurden folgende Verfahren angewandt:

2.9.1 Vier-Felder-Tafeln

Die Auswertung der kategorialen Daten des Patientenkollektivs wurden in den in der Statistik üblichen Vier-Felder-Tafeln (Tab. 2) dargestellt (Hilgers, 2007). Anhand dieser wurden statistische Maßzahlen ermittelt, die zur Einschätzung der Treffsicherheit der konventionellen Zytologie, der manuellen DNA-Bildzytometrie und der automatischen DNA-Bildzytometrie im Vergleich zum operativen histologischen Befund als Referenzstandard dienen.

Erkrankung	Test		
	positiv	negativ	gesamt
positiv	a	b	a+b
negativ	c	d	c+d
gesamt	a+c	a+d	a+b+c+d

Tabelle 2: Prinzip der Vier-Felder-Tafel

Berechnungen der Treffsicherheit werden unter Verwendung folgender Parameter vorgenommen:

Sensitivität: Die Sensitivität errechnet sich als Quotient aus der Anzahl der richtig-positiv getesteten Erkrankten und der Anzahl aller Erkrankten. Sie beschreibt somit den Anteil der Patienten, bei denen durch ein positives Testergebnis die Diagnose einer Erkrankung richtig gestellt wird.

$$\text{Sensitivität} = \left(\frac{a}{a+b} \right)$$

Spezifität: Die Spezifität errechnet sich genau "spiegelverkehrt" zur Sensitivität als Quotient aus der Anzahl der richtig-negativ getesteten Gesunden und der Anzahl aller Gesunden und beschreibt somit den Anteil der Patienten, bei denen durch ein negatives Testergebnis die Diagnose einer Erkrankung ausgeschlossen wird.

$$\text{Spezifität} = \left(\frac{d}{c+d} \right)$$

Die **Falsch-Positiv-Rate** beschreibt die Rate von Fällen, die im Test fälschlich als positiv klassifiziert wurden, die aber in Wirklichkeit negativ, also gesund sind. Die Falsch-Positivrate ergänzt sich mit der jeweiligen Spezifität zu 100%.

$$\text{Falsch-Positiv-Rate} = \left(\frac{c}{c+d} \right)$$

Die **Falsch-Negativ-Rate** beschreibt die Rate von Fällen, die fälschlich als negativ klassifiziert wurden, die aber in Wirklichkeit positiv, also erkrankt sind und ist somit der prozentuale Wert von übersehenen Krankheitsdiagnosen. Sie ergänzt sich mit der Sensitivität zu 100%.

$$\text{Falsch-Negativ-Rate} = \left(\frac{b}{a+b} \right)$$

Der **positive Prädiktionswert** (PPV, *positive predictive value*) errechnet sich als Quotient aus der Anzahl der im Test richtig-positiven Fälle und der Anzahl aller als positiv erkannten Fälle.

$$\text{PPV} = \left(\frac{a}{a+c} \right)$$

Der **negative Prädiktionswert** (NPV, *negative predictive value*) errechnet sich komplementär hierzu als Quotient aus der Anzahl der richtig-negativen Fälle und der Anzahl aller als negativ erkannten Fälle.

$$\text{NPV} = \left(\frac{b}{b+d} \right)$$

Die **Gesamttreffericherheit** schließlich beschreibt den Prozentsatz aller richtig positiven und richtig negativen erkannten Diagnosen, bezogen auf die Gesamtzahl aller Fälle.

$$\text{Gesamttreffericherheit} = \left(\frac{a+d}{a+b+c+d} \right)$$

Zur statistischen Auswertung von Vier-Felder-Tafeln wurde der Chi²-Test angewandt; als Signifikanzniveau wurde, wie international üblich, ein α -Fehler von $< 5\%$, entsprechend einem p-Wert von $p < 0,05$, definiert.

2.9.2 Cohens Kappa

Zur Berechnung der Übereinstimmungen zwischen zwei Beurteilungen für kategoriale Daten wurde der Kappa-Koeffizient (κ) nach Cohen verwendet (Cohen 1960). Er wird berechnet als Quotient aus der Differenz zwischen beobachteter (p_0) und zufallsbedingt zu erwartender Übereinstimmung (p_e) und der Differenz aus maximal möglicher (1, entspricht 100%) und zufallsbedingt zu erwartender Übereinstimmung.

Die Formel lautet dementsprechend:

$$\kappa = \frac{p_0 - p_e}{1 - p_e}$$

Kappa kann Werte zwischen 0 und 1 annehmen. Zur Beurteilung wurde die folgende Einteilung (mit englischsprachiger Entsprechung in Klammern) gewählt (Grouven et al., 2007):

- Kappa < 0,20 schwache Übereinstimmung ("*poor agreement*")
- Kappa 0,21-0,40 leichte Übereinstimmung ("*fair agreement*")
- Kappa 0,41-0,60 mittelmäßige Übereinstimmung ("*moderate agreement*")
- Kappa 0,61-0,80= gute Übereinstimmung ("*substantial agreement*")
- Kappa >0,81 sehr gute Übereinstimmung ("*almost perfect agreement*")

2.9.3 Bland-Altman-Diagramme

Bland-Altman-Diagramme wurden in den 80er Jahren in die Statistik und Biometrie eingeführt (Bland & Altman, 1986). In ihnen werden in einem kartesischen Koordinatensystem die Mittelwerte zweier Messungen auf der x-Achse und die Differenzen zwischen den beiden Messwerten auf der y-Achse aufgetragen. Ferner wird die Linie eingezeichnet, die dem Mittelwert der Differenzen der Messwerte entspricht, ergänzt um zwei Linien, die den zuvor festgelegten Vertrauensbereich der Messwerte repräsentieren (z.B. Mittelwert der Differenzen $\pm$ SD oder $\pm$ 2 SD etc.), der möglichst eingehalten werden soll.

Die Diagramme erleichtern die Erkennung von systematischen Messfehlern, von Schwankungskorridoren und von Abhängigkeiten errechneter Messwertdifferenzen von der Höhe der Messwerte. Insofern bieten sie Informationen, die über die Parameter der deskriptiven Statistik deutlich hinausgehen.

Zur Testung von statistischer Signifikanz wurde bei der Auswertung der Bland-Altman-Diagramme der Spearman'sche Rangkorrelationskoeffizient (r_s) herangezogen, wobei ein p-Niveau von $p < 0,05$ als Beleg für Signifikanz gewertet wurde.

2.10 Ethikvotum

Mit Datum vom 30.1.2015 liegt zu der hier vorgestellten Studie ein positives Votum der Ethikkommission der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf vor ("Methodische Untersuchungen zur Automatisierung und Variablenoptimierung der DNA-Bildzytometrie", Studiennummer 4942).

3 Ergebnisse

3.1 Patientenkollektiv

Für die vorliegende Studie wurden 92 Feinnadelpunktate der Schilddrüse von 92 verschiedenen Patienten herangezogen, von denen bereits ein Messergebnis der manuellen DNA-Bildzytometrie im Rahmen der zytologischen Routinediagnostik vorlag.

Von 46 dieser Patienten (50%) konnte vor Beginn der automatischen Messreihe ein histopathologisches und klinisches Follow-up ermittelt werden. Diese 46 Präparate wurden im Nachhinein mit dem automatischen DNA-bildzytometrischen Messsystem erneut untersucht. Letztendlich konnten für 40 der 46 Präparate (87,0%) Messergebnisse mit dem automatischen DNA-Zytometer generiert und für die weitere Auswertung herangezogen werden.

3.1.1 Ausgeschlossene Präparate

Sechs Präparate mussten von der automatischen DNA-Zytometrie ausgeschlossen werden.

In zwei Fällen war die Färbung über die Jahre trotz dunkler Lagerung zu stark abgeblasst (Abb. 18, links), in vier Fällen ergaben sich Detektionsprobleme aufgrund starker blutiger Überlagerung und zu dicht angeordneter und sich überlappender Analysezellkerne (Abb. 18, rechts).

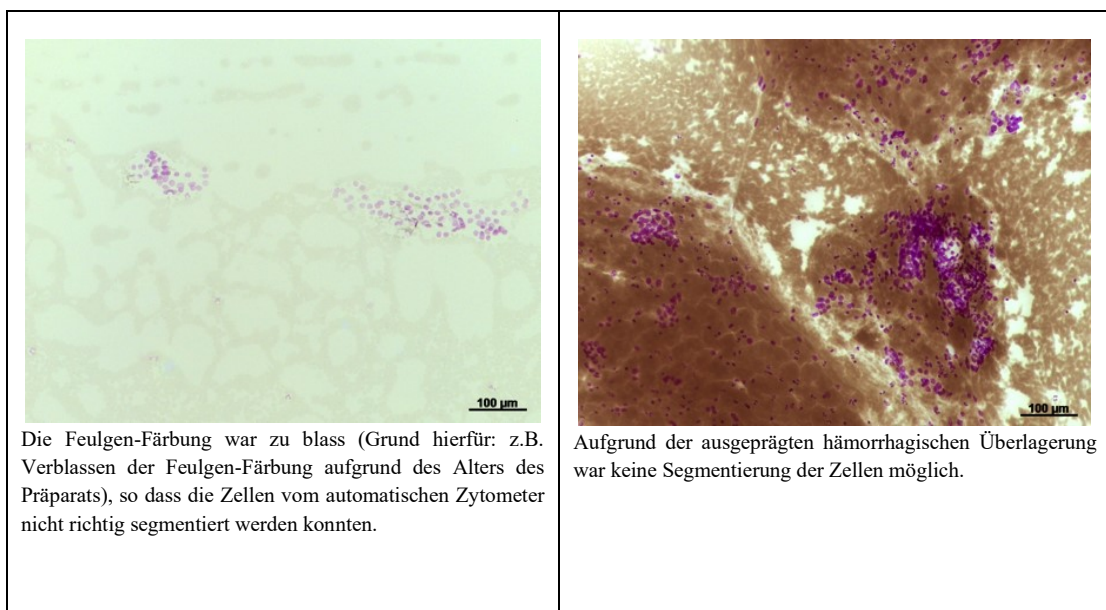


Abbildung 18: Zwei repräsentative Beispiele von der Messung ausgeschlossener Präparate. Links: zu starke Abblässung nach mehrjähriger Lagerung; rechts: zu starke Hämorrhagie und Zellüberlappungen.

3.1.2 Alters- und Geschlechterverteilung

Insgesamt wurden 38 Frauen (95,0%) und zwei Männer (5,0%) untersucht. Die Altersverteilung der Patienten zum Zeitpunkt der Diagnose wird in Abbildung 19 dargestellt. Es zeigt sich eine Altersverteilung von 21-80 Jahren, relativ gleichmäßig über alle Altersklassen verteilt, mit einem Maximum in der Altersgruppe 51-60. Das mittlere Alter aller Patienten zum Zeitpunkt der Diagnose beträgt 52,9 +/- 15,4 [SD] Jahre, das der Frauen 52,6 +/- 14,9 [SD] Jahre, das der Männer bei 59,5 +/- 20,5 [SD] Jahre.

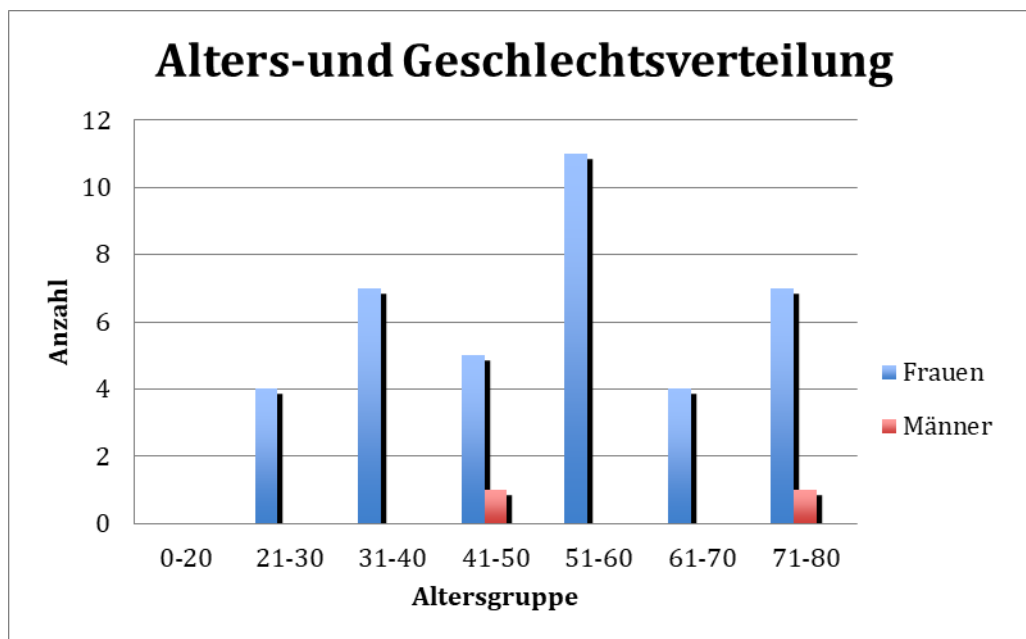


Abbildung 19: Alters- und Geschlechterverteilung im Patientenkollektiv (n=40)

3.2 Morphologische Diagnostik

3.2.1 Zytologie

In allen 40 Fällen lagen diagnostisch auswertbare Feinnadelpunktate vor; dabei waren sieben Fälle negativ, 25 zweifelhaft, sechs dringend verdächtig und zwei positiv (Tabelle 3, Abbildung 20).

Zytologische diagnostische Kategorie	Zahl der Fälle
negativ	7
zweifelhaft	25
dringender Verdacht	6
positiv	2

Tabelle 3: Anteile der einzelnen zytologischen diagnostischen Kategorien an den untersuchten 40 Präparaten

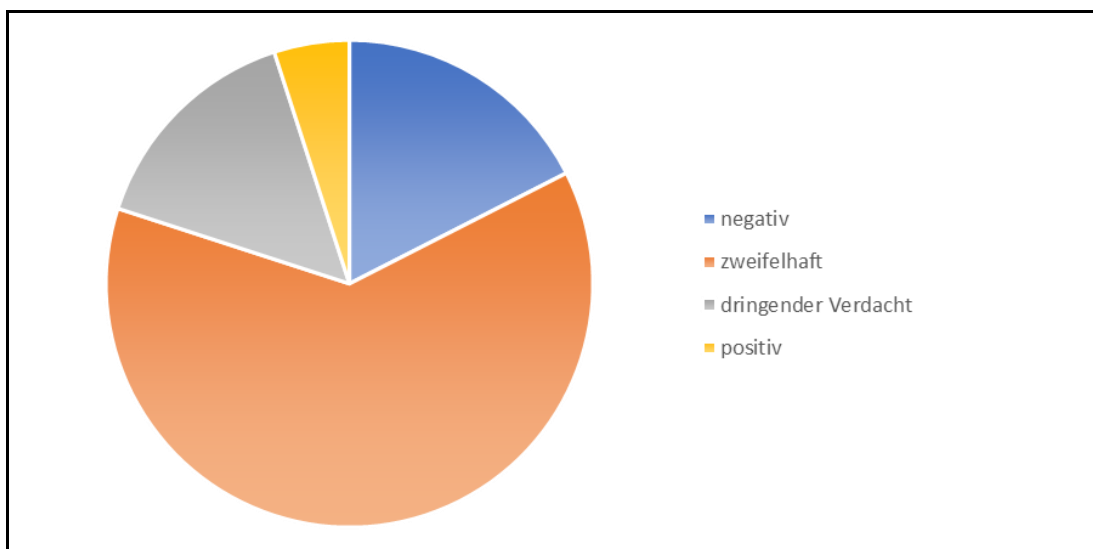


Abbildung 20: Kreisdiagramm zur Verteilung der zytologischen Diagnosekategorien

In Tabelle 4 und Abbildung 21 sind die textlichen Diagnosen der konventionellen Zytologie zu den 45 Fällen wiedergegeben:

Zytologische Klartextdiagnosen	Zahl der Fälle
Thyreoiditis	2
Thyreoiditis mit ausgeprägter follikulärer Hyperplasie	4
Follikuläre Hyperplasie	17
Kolloidstruma	7
Neoplastische Transformation	8
Noduläre Hyperplasie	1
Follikuläres Adenom	1

Tabelle 4: Zytologische Klartextdiagnosen

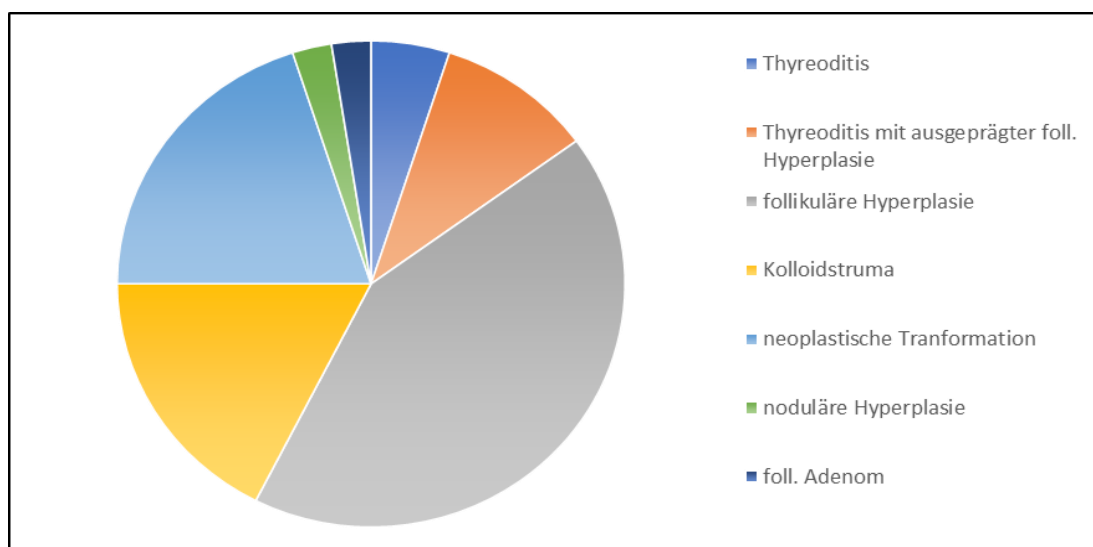


Abbildung 21: Kreisdiagramm zur Darstellung der zytologischen Klartextdiagnosen

3.2.2 Histologie

Von allen 40 Fällen lag eine histologische Diagnose vor (Tab. 5 und Abb. 22).

Histologische diagnostische Kategorie	Zahl der Fälle
reaktiv / Struma	15
Adenom	21
Karzinom	4

Tabelle 5: Anteile der einzelnen histologischen diagnostischen Kategorien der 40 Fälle

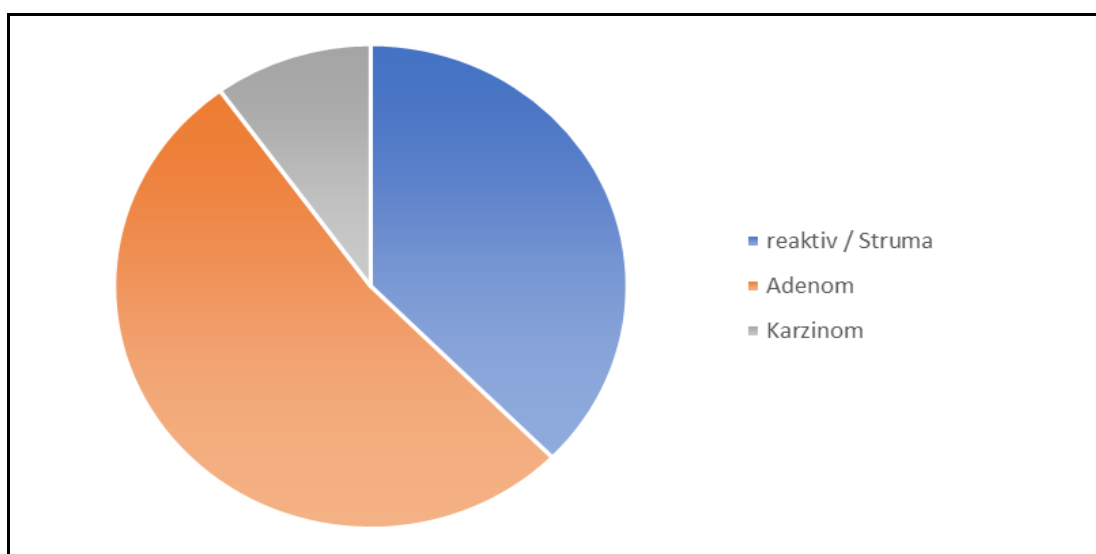


Abbildung 22: Kreisdiagramm zur Darstellung der histologischen diagnostischen Kategorien

3.2.3 Treffsicherheit der konventionellen Zytologie

Um die Treffsicherheit der zytopathologischen Diagnosen auswerten zu können, ist eine Dichotomisierung der vier Stufen der Malignitätswahrscheinlichkeit notwendig. Diese Aufteilung dient der Errechnung von Sensitivität und Spezifität. Voraussetzung ist, dass der Test eine binäre Auswahl in „Test-positiv“ oder „Test-negativ“ bietet.

Bei den zweifelhaften und dringend verdächtigen Präparaten musste demnach entschieden werden, ob sie als „Test-positiv“ oder „Test-negativ“ eingestuft werden sollten. Da im Falle der Schilddrüse durch die Zytologie primär geklärt werden soll, ob in einem ehemaligen Strumaendemiegebiet wie Deutschland eine OP-Indikation zu stellen ist und somit auch zweifelhafte und dringend verdächtige Fälle der gleichen Primärtherapie wie die positiven Fälle zugeführt werden sollen, wurden diese drei Gruppen hier als "positiv" zusammengefasst und den zytologisch negativen Fällen gegenübergestellt. "Positiv" ist hier also gleichbedeutend mit der Diagnose oder der Verdachtsdiagnose eines malignen Tumors oder einer Neoplasie (maligne oder benigne).

Die Vierfeldertafel zeigt einen Überblick über die absolute Häufigkeit der Ergebnisse für alle 40 Fälle (Tab. 6). Dabei ließ sich nicht mehr eruieren, ob für die zytologische Untersuchung Zellen aus derjenigen Läsion, an der die gravierendste histologische Veränderung gesehen wurde, oder aus einem anderen Abschnitt des Organs erfasst waren. Eine solche Korrelation ließ sich retrospektiv nicht mehr herstellen.

Zytologische Diagnose	Histologische Diagnose	
	Tumor (gutartig / bösartig)	Struma / reaktiv
negativ	3	3
zweifelhaft / dringender Verdacht / positiv	17	17

Tabelle 6: Vierfeldertafel zur diagnostischen Treffsicherheit der Zytologie aller 40 Fälle

Die entsprechenden Kennwerte betragen für die Sensitivität 85,0%, für die Spezifität 15,0%, für den positiven prädiktiven Vorhersagewert (PPV) 50,0%, für den negativen prädiktiven Vorhersagewert (NPV) 50,0% und für die Gesamttreffsicherheit 50,0%.

3.3 Vergleichende Statistik zur manuellen und automatischen DNA-Bildzytometrie

3.3.1 Zahl der Analyse- und Referenzzellen

Mit dem automatischen Messsystem wurden deutlich mehr Analyse- und Referenzzellen erfasst als bei der manuellen Messung (Tab. 7), bei der von vornherein die Zahl zu erfassender Zellen begrenzt war (Ziel: ca. 30 Referenz- und ca. 300 Analysezellen). Die Mittelwerte lagen um den Faktor 7,9 bzw. 10,2 höher. Die enorme Spannweite der Werte sowie die hohen Werte für die Standardabweichung und den Variationskoeffizienten bei den automatischen Messungen reflektieren dabei die unterschiedliche Zellularität, die Schilddrüsenpunktate aufweisen können.

Parameter	Anzahl der Analysezellen		Anzahl der Referenzzellen	
	Manuelle Messung	Automatische Messung	Manuelle Messung	Automatische Messung
Minimum	154	228	30	28
Maximum	426	14333	46	1797
Mittelwert	314,63	2474,68	36,18	370,38
Standardabweichung	59,72	3178,82	4,82	489,54
Variationskoeffizient	0,19	1,28	0,13	1,32

Tabelle 7: Statistische Kennwerte für die Zahl der gemessenen Analysezellen und Referenzzellen

3.3.2 Mittlere Referenzzell-IOD

Im Vergleich der mittleren IOD der Referenzzellen, die fallbezogen definitionsgemäß als 2c gesetzt wird, zeigten sich auf der Basis der deskriptiven Statistik gut vergleichbare Werte (Tab. 8).

Parameter	Mittlere Referenzzell-IOD	
	Manuelle Messung	Automatische Messung
Minimum	219,99	218,88
Maximum	492,18	469,96
Mittelwert	324,27	305,58
Standardabweichung	79,45	53,22
Variationskoeffizient	0,25	0,17

Tabelle 8: Statistische Kennwerte für die mittlere Referenzzell-IOD der manuellen und der automatischen Messungen; die Angaben beziehen sich auf nur 39 statt auf 40 Fallvergleiche, da ein Wert bei der manuellen Messung irrtümlich nicht abgespeichert worden war.

3.3.3 Minimaler C-Wert

Bei der manuellen Messung befand sich der minimale C-Wert zwischen 1,28c und 2,50c. Bei der automatischen Messung lag der kleinste minimale C-Wert bei 1,23c, der höchste minimale C-Wert bei 2,50c. Die entsprechende Datenübersicht ist in Tab. 9 wiedergegeben. In den Abb. 23 und 24 sind das zugehörige Streudiagramm gezeigt und ein Bland-Altman-Diagramm dargestellt.

Parameter	Minimaler C-Wert	
	Manuelle Messung	Automatische Messung
Minimum	1,28	1,23
Maximum	2,50	2,50
Mittelwert	1,76	1,70
Standardabweichung	0,23	0,31
Variationskoeffizient	0,13	0,18

Tabelle 9: Statistische Kennwerte für den minimalen C-Wert der manuellen und der automatischen Messungen

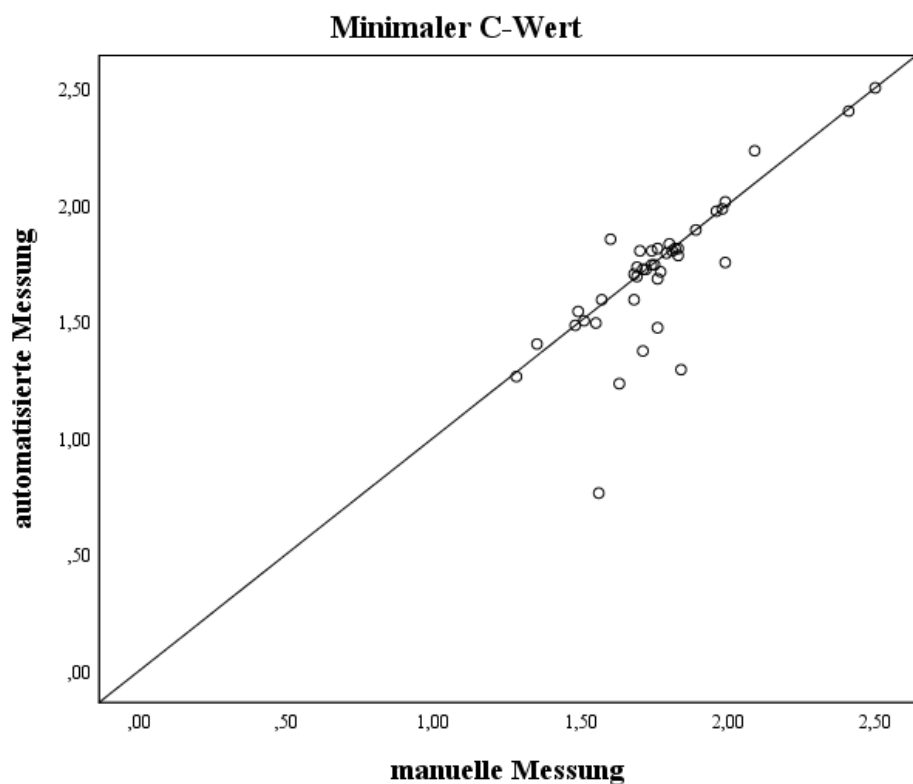


Abbildung 23: Streudiagramm der Messwerte der minimalen C-Werte

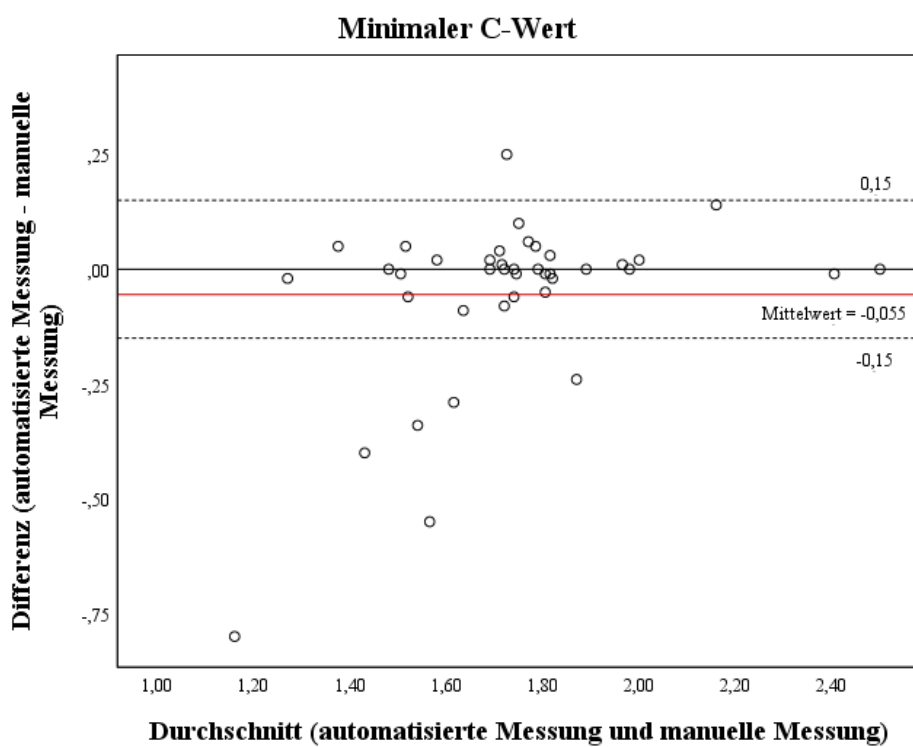


Abbildung 24: Bland-Altman-Diagramm der Messwerte der minimalen C-Werte

In dem Streudiagramm in Abbildung 23 ist dargestellt, dass die minimalen C-Werte der automatischen und manuellen Messungen größtenteils gut übereinstimmen, da sie auf bzw. nahe an der winkelhalbierenden Diagonalen liegen. Eine leichte Verschiebung einiger Messpaare zu Werten unterhalb des Vertrauensbereiches war zu erwarten, da bei der manuellen Messung blässere als der „optische Durchschnitt“ angefärbte Zellen im Gegensatz zur automatischen Messung meist nicht mitgemessen werden konnten.

Der Bland-Altman-Plot in Abbildung 24 weist aus, dass 82,5% der Differenzen von automatischen und manuellen Messungen des minimalen C-Wertes innerhalb des vorgegebenen Vertrauensbereiches von $\pm 0,15c$ liegen. Außerdem liegt der Mittelwert der Differenzen beider Messergebnisse mit $-0,055$ annähernd beim Idealwert von Null. Außerhalb der Vertrauensbereiches liegen insgesamt sechs Messwertpaare. Das 95%-Konfidenzintervall beträgt 57%-94,2%. Es ergab sich keine Abhängigkeit von Differenzen und Durchschnittswerten der Messpaare der minimalen C-Werte (Korrelationskoeffizient nach Spearman: $p = 0,107$).

3.3.4 Maximaler C-Wert

Bei der manuellen Messung bewegte sich der maximale C-Wert zwischen 3,71c bis maximal 17,49c, bei der automatischen Messung zwischen 3,75c und 17,60c.

Die entsprechende Datenübersicht ist in Tab. 10 wiedergegeben. In den Abb. 25 und 26 sind das zugehörige Streudiagramm und ein Bland-Altman-Diagramm dargestellt.

Parameter	Maximaler C-Wert	
	Manuelle Messung	Automatische Messung
Minimum	3,71	3,75
Maximum	17,49	17,60
Mittelwert	8,31	7,89
Standardabweichung	3,72	3,50
Variationskoeffizient	0,45	0,45

Tabelle 10: Statistische Kennwerte für den maximalen C-Wert der manuellen und der automatischen Messungen

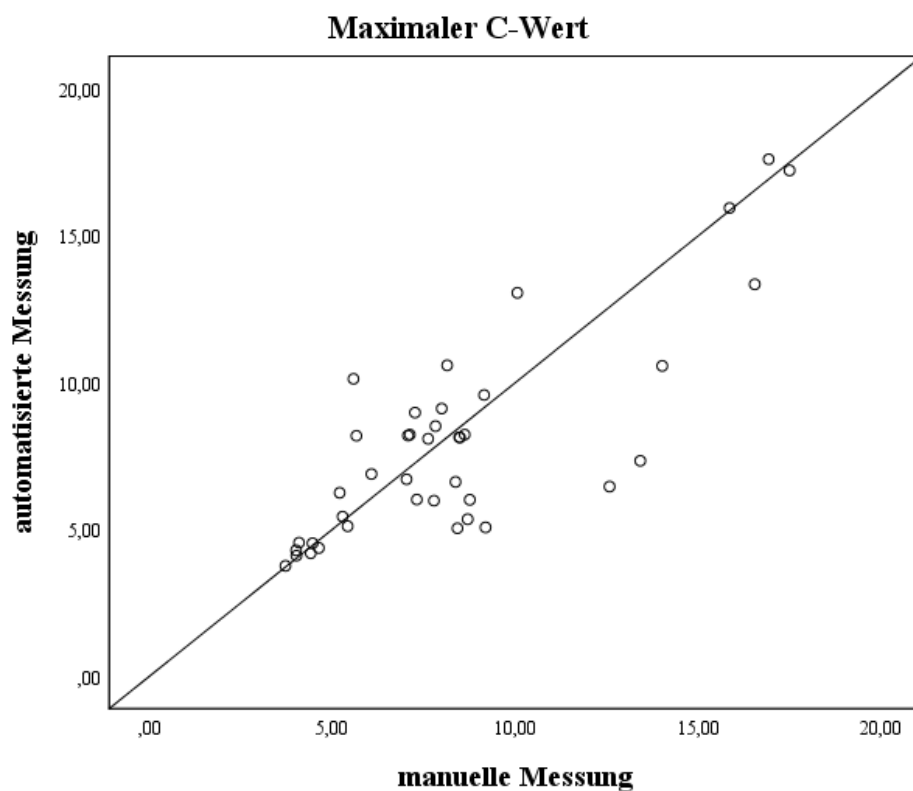


Abbildung 25: Streudiagramm der Messwerte der maximalen C-Werte

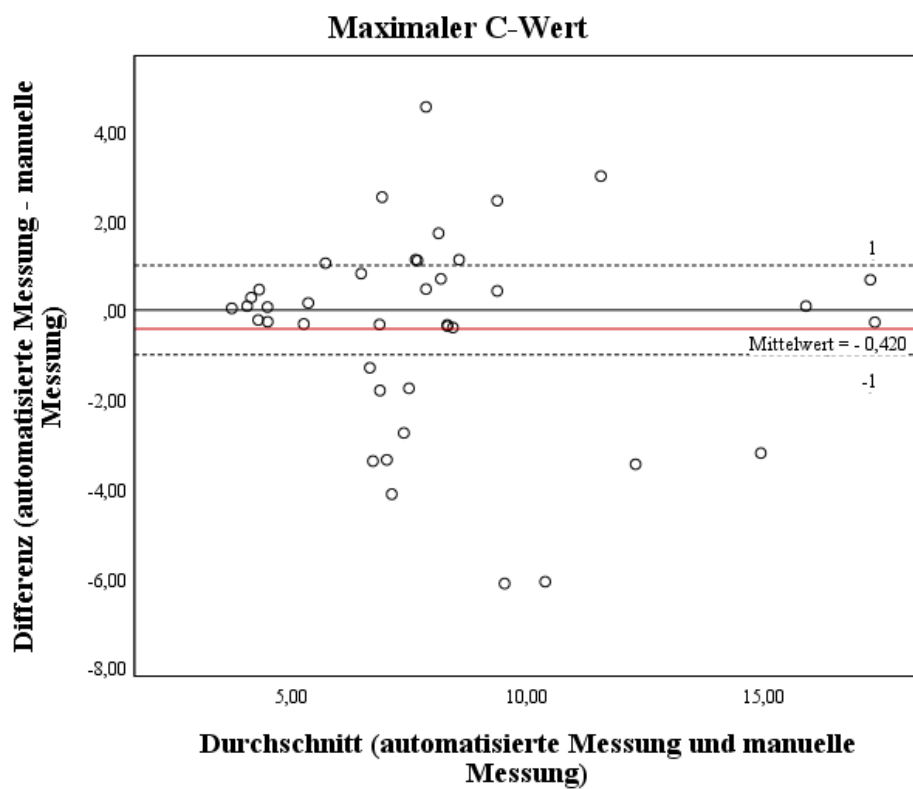


Abbildung 26: Bland-Altman-Diagramm der Messwerte der maximalen C-Werte

Das Streudiagramm in Abbildung 25 zeigt durchaus eine Lage der Datenpunkte in gutem Bezug zur winkelhalbierenden Diagonalen, aber keine sehr enge Korrelation der Messwerte zwischen den automatischen und den manuellen Messungen. Bei Ausreißern überwiegen höhere Werte bei der manuellen Messung, möglicherweise durch eine selektive Suche nach größeren und dunkler angefärbten Zellkernen zu erklären.

Abbildung 26 zeigt für den maximalen C-Wert, dass 50% der Differenzen von automatischen und manuellen Messungen des maximalen C-Wertes innerhalb eines (willkürlich festgelegten) Intervalls von $\pm 1,00c$ liegen. Das 95% Konfidenzintervall beträgt 17,4%-74,6%. Es gibt keine Abhängigkeit von Differenzen und Durchschnittswerten der Messpaare, was mit dem Korrelationskoeffizienten nach Spearman ($p = 0,835$) ermittelt wurde.

Die mittlere Differenz von $-0,420c$ belegt, dass manuelle Messungen zumindest im Trend einen höheren Maximalwert aufweisen. Andererseits sind deutliche Unterschiede zwischen $+4c$ und $-6c$ bei einzelnen Fällen möglich.

Zusammenfassend sind die Maximalwerte zwischen den beiden Messreihen zwar recht gut vergleichbar, können aber im Einzelfall unerwarteterweise doch recht weit voneinander abweichen.

3.3.5 5c EE / 5cER

Die Messwerte für den 5cEE lagen bei der manuellen Messung zwischen 0 und 51, bei der automatischen Messung zwischen 0 und 181. Die entsprechenden Daten sind in Tab. 11 wiedergegeben. In den Abb. 27 und 28 sind das zugehörige Streudiagramm und ein Bland-Altman-Diagramm dargestellt.

Parameter	5cEE	
	Manuelle Messung	Automatische Messung
Minimum	0	0
Maximum	51	181
Mittelwert	10,23	15,20
Standardabweichung	13,62	31,25
Variationskoeffizient	1,33	2,06

Tabelle 11: Statistische Kennwerte für den 5cEE der manuellen und der automatischen Messungen

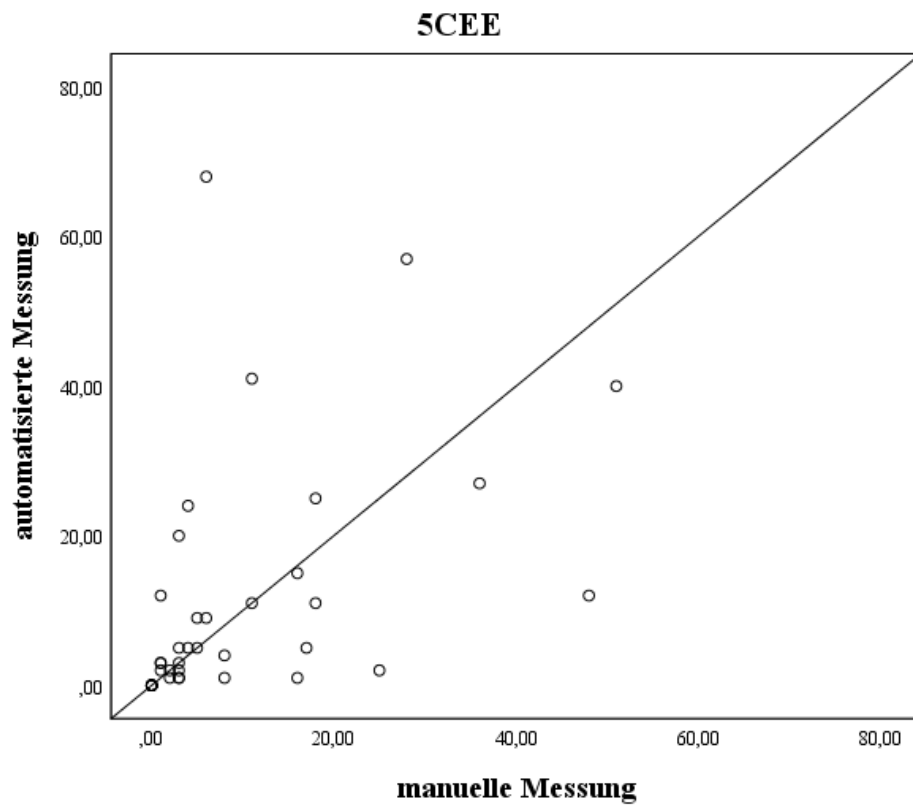


Abbildung 27: Streudiagramm der Messwerte für den 5cEE (39 Messwertpaare, siehe Text)

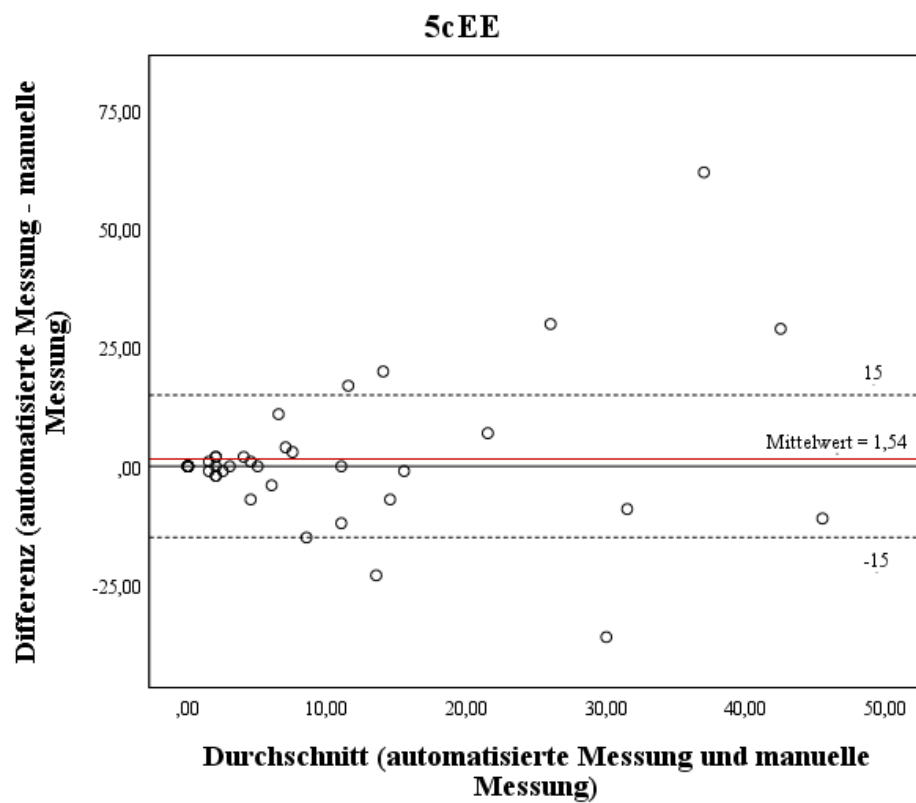


Abbildung 28: Bland-Altman-Diagramm für den 5cEE (39 Messwertpaare, siehe Text)

Das Streudiagramm in Abbildung 27 zeigt, dass bei Auftreten von nur wenigen 5cEE die manuellen und die automatischen Messergebnisse gut korrelieren ($r_s = 0,542$, $p < 0,001$), aber bei einer höheren Zahl von 5cEE eher zur Divergenz neigen. 82% der Differenzen von automatischen und manuellen Messungen der 5cEE liegen gemäß dem Bland-Altman-Diagramm in Abb. 28 innerhalb des (willkürlich festgelegten) Intervalls von ± 15 liegen. Das 95% Konfidenzintervall beträgt 53,3%-93,7%.

Es lag keine statistisch signifikante Korrelation von Differenzen und Durchschnittswerten der Messpaare vor. Die mittlere Differenz der Messergebnisse von 9,425 belegt dabei noch einmal, dass im Durchschnitt die automatischen Messungen zu einer Detektion von mehr 5cEEs führten als die manuellen Messungen. Zur verlässlicheren Bewertung von Tendenzen in den Datenverteilungen wurde ein Ausreißer (5cEE manuell 42, automatisch 181) nicht berücksichtigt, weswegen sich die Diagramme nur auf 39 Fälle beziehen.

Die Messwerte für den 5cER, den Prozentwert der 5cEE bezogen auf die Zahl der Analysezellen, bewegte sich bei der manuellen Messung zwischen 0% und 26,82% und bei der automatischen Messung zwischen 0% und 7,33%. Die entsprechenden Daten zeigt Tab. 12. In den Abb. 29 und 30 sind das zugehörige Streudiagramm und ein Bland-Altman-Diagramm wiedergegeben.

Parameter	5cER	
	Manuelle Messung	Automatische Messung
Minimum	0%	0%
Maximum	26,82%	7,33%
Mittelwert	3,54%	0,90%
Standardabweichung	5,33%	1,41%
Variationskoeffizient	1,50	1,57

Tabelle 12: Statistische Kennwerte für den 5cER der manuellen und der automatischen Messungen

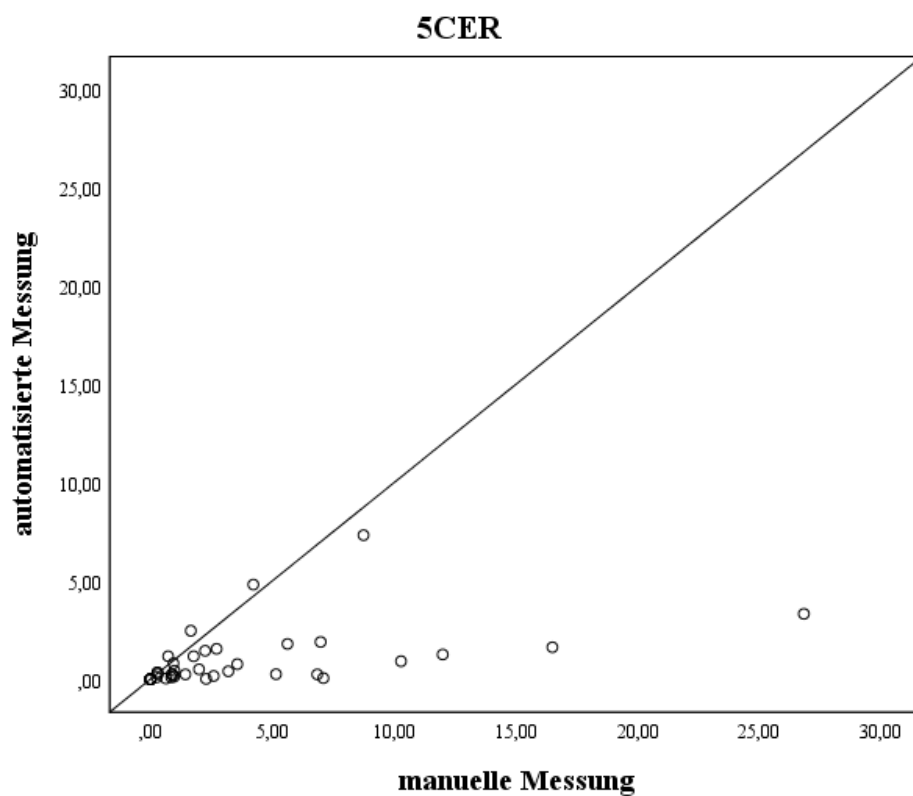


Abbildung 29: Streudiagramm der Messwerte für den 5cER

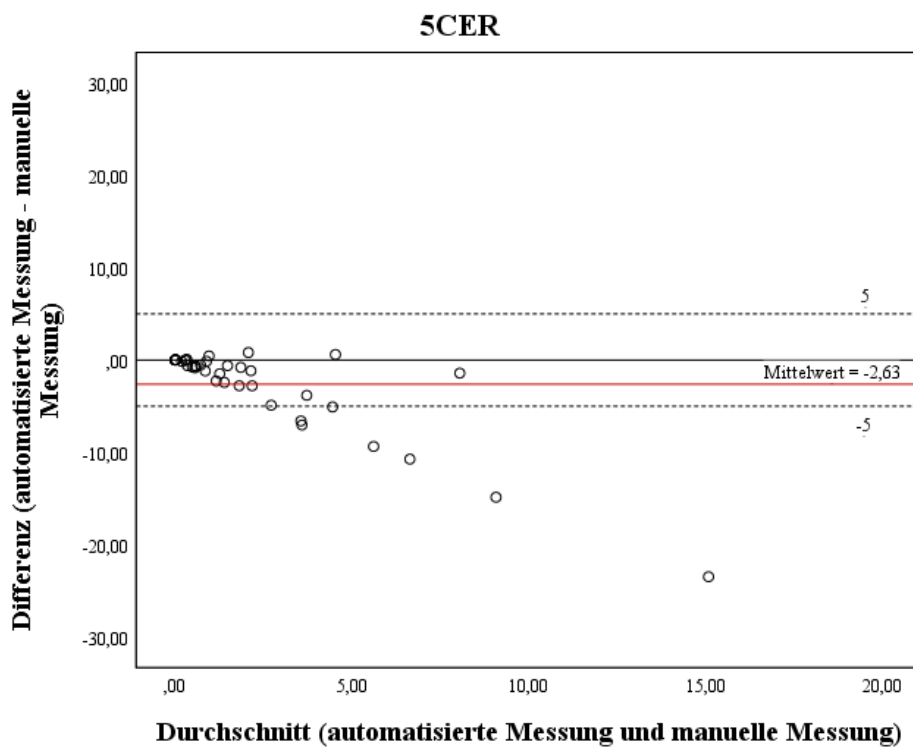


Abbildung 30: Bland-Altman-Diagramm der Messwerte für den 5cER

Das Streudiagramm in Abbildung 29 zeigt noch einmal anschaulich, dass die 5cER-Werte für die automatische Messung in aller Regel klein sind und meistens 5% nicht überschreiten. So lange auch für die manuelle Messung ein Wert bis 5% vorliegt, ist eine Korrelation beider Messreihen gegeben ($r_s = 0,481$, $p = 0,002$). Daneben kommen aber einige Fälle vor, in denen der manuelle 5cER deutlich höher liegt, während der automatische 5cER niedrig bleibt.

Genau diese Fälle sind es, die dann im Bland-Altman-Diagramm (Abb. 30) unter dem willkürlich festgelegten Vertrauensbereich von $\pm 5\%$ liegen. In diesem wiederum liegen auch für den 5cER ca. 80% der Messwertdifferenzen. Das 95% Konfidenzintervall beträgt 54%-92,6%. Es ist mittels des Korrelationskoeffizienten nach Spearman ($p < 0,001$) eine statistisch signifikante Beziehung von Differenzen und Durchschnittswerten der Messwertpaare zu beobachten. Diese Abhängigkeit beeinflusst die nicht parametrische Statistik bei willkürlich festgelegten Grenzen nicht.

Fasst man die Ergebnisse von 5cEE und 5cER vergleichend zusammen, so sind die Korrelationen für den 5cER wie erwartet höher.

3.3.6 9c EE / 9cER

Die Datenverteilung für den 9cEE für die manuellen und die automatischen Messungen zeigt Tabelle 13. Die Messwerte für den 9cEE bewegten sich bei der manuellen Messung zwischen 0 und 4, bei der automatischen Messung zwischen 0 und 18 mit Mittelwerten von 0,53 bzw. 1,05.

Parameter	9cEE	
	Manuelle Messung	Automatische Messung
Minimum	0	0
Maximum	4	18
Mittelwert	0,53	1,05
Standardabweichung	1,15	3,19
Variationskoeffizient	2,20	3,04

Tabelle 13: Statistische Kennwerte für den 9cEE der manuellen und der automatischen Messungen

Beiden Messreihen ist gleich, dass nur in einem kleineren Teil der Fälle überhaupt 9cEEs auftraten. Bei beiden Messreihen wurden in 10 von 40 Fällen 9cEEs gefunden (25%). In Abbildung 31 ist die Häufigkeitsverteilung dargestellt.

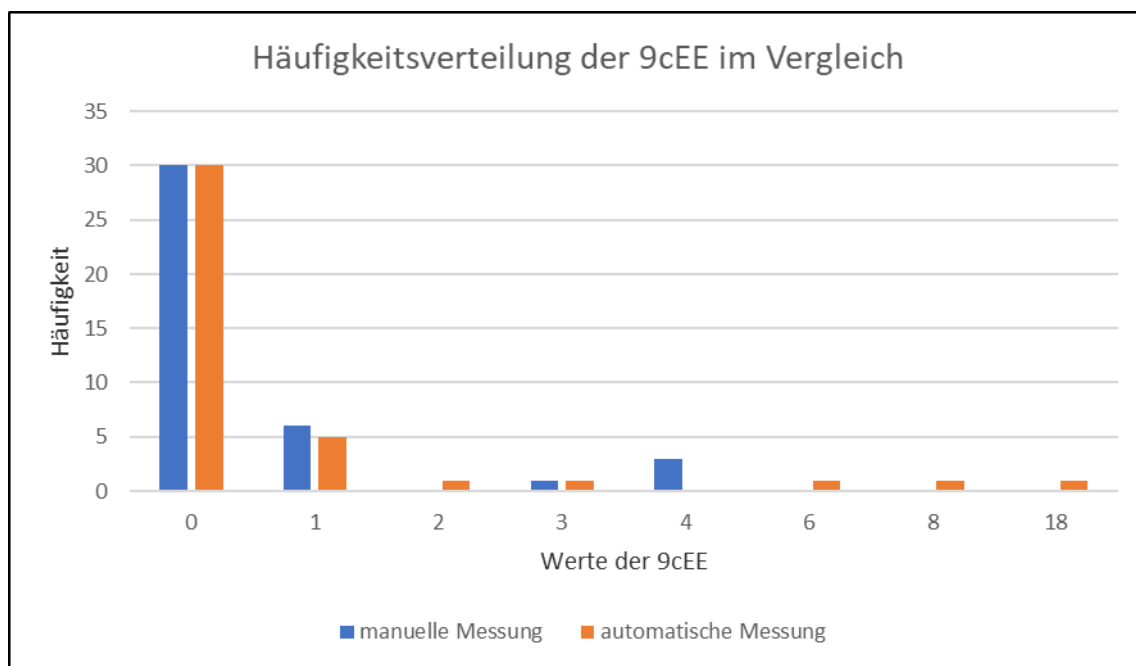


Abbildung 31: Häufigkeitsverteilung der 9cEE im Vergleich manueller und automatischer Messungen

Tabelle 14 gibt die entsprechende Datenverteilung der Messwertpaare tabellarisch wieder. Insgesamt ist die Anzahl der gefundenen 9cEE bei beiden Messmethoden eher gering. Mit der automatischen Messung konnten einige wenige Fälle mit 9cEE > 4 erkannt werden. Andererseits konnte in drei Fällen der singuläre 9cEE der manuellen Messung automatisch nicht identifiziert werden.

Parameter	9cEE (automatische Messung)				
	0	1	2	≥ 3	gesamt
9cEE (manuelle Messung)					
0	27	3	0	0	30
1	3	1	1	1	6
2	0	0	0	0	0
≥ 3	0	1	0	3	4
gesamt	30	5	1	4	40

Tabelle 14: Übereinstimmungsmatrix für die Werte des 9cEE bei manueller bzw. automatischer Messung; Chi²-Test: $p < 0,0001$

Für die 40 Fälle ergibt sich ohne weitere Subgruppierung, also bezogen auf jeweils vier Gruppen, eine Gesamtübereinstimmung von 77,5% (31/40) mit einem Kappa Wert von 0,45 (moderate agreement). Ergänzend wurden die Übereinstimmungen der Messergebnisse für verschiedene Grenzziehungen untersucht. Bei einer Unterscheidung von

zwei Gruppen für $9cEE = 0$ gegen $9cEE > 0$ ergab sich eine Übereinstimmung von 85% (34/40) und ein Kappa-Wert von 0,60 (moderate agreement). Für einen Vergleich in drei Gruppen ($9cEE = 0$, $0 < 9cEE < 3$, $9cEE \geq 3$) liegt die Übereinstimmung bei 80% (32/40) und der Kappa-Wert bei $= 0,506$ (moderate agreement).

Tabelle 15 zeigt, dass sich die Fallzahlen mit höheren Werten bei der einen oder der anderen Methode mit einem Verhältnis von 5:4 fast die Waage halten.

9cEE-Wert	Gesamtzahl der Fälle (n = 40)	Fälle mit 9cEE > 0 (n = 13)
manuell > automatisch	4	4
manuell = automatisch	31	4
manuell < automatisch	5	5

Tabelle 15: Einteilung der Fälle mit 9cEE nach ihrer Häufigkeit bei beiden Messungen, n=40 (alle Fälle) bzw. n = 13 (Fälle mit wenigstens in einer der beiden Messungen $9cEE > 0$)

Die Messwertpaare für den 9cEE sind ergänzend in der Abb. 32 und 33 als Streudiagramm und in Form eines Bland-Altman-Diagramms wiedergegeben.

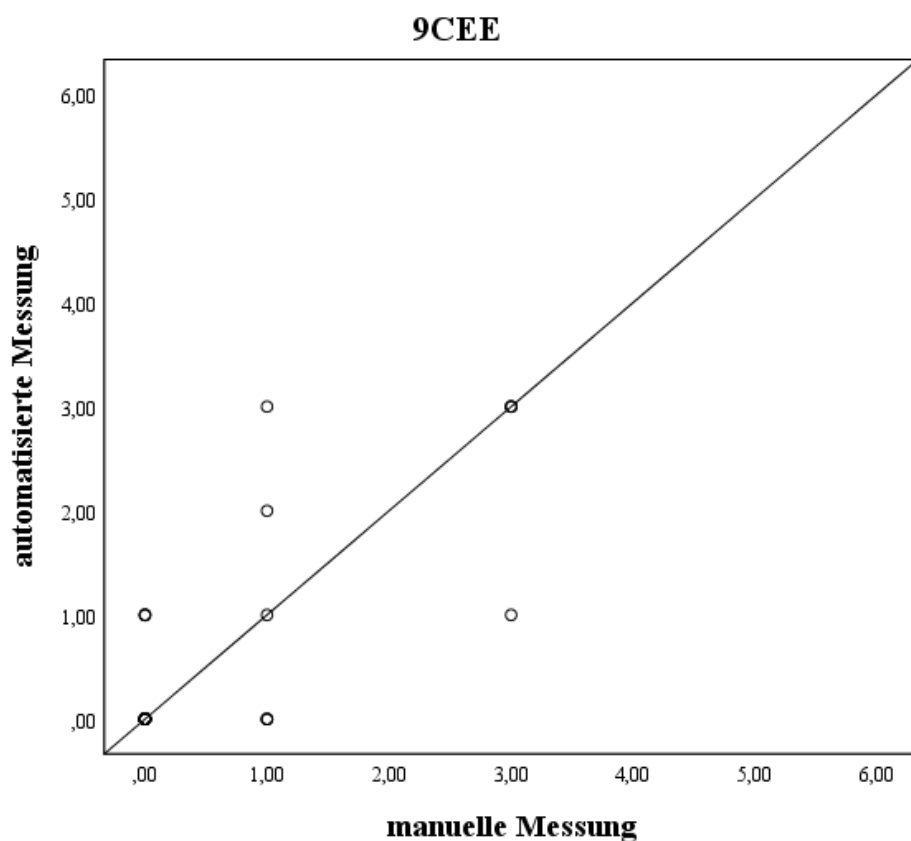


Abbildung 32: Streudiagramm der Messwerte für den 9cEE

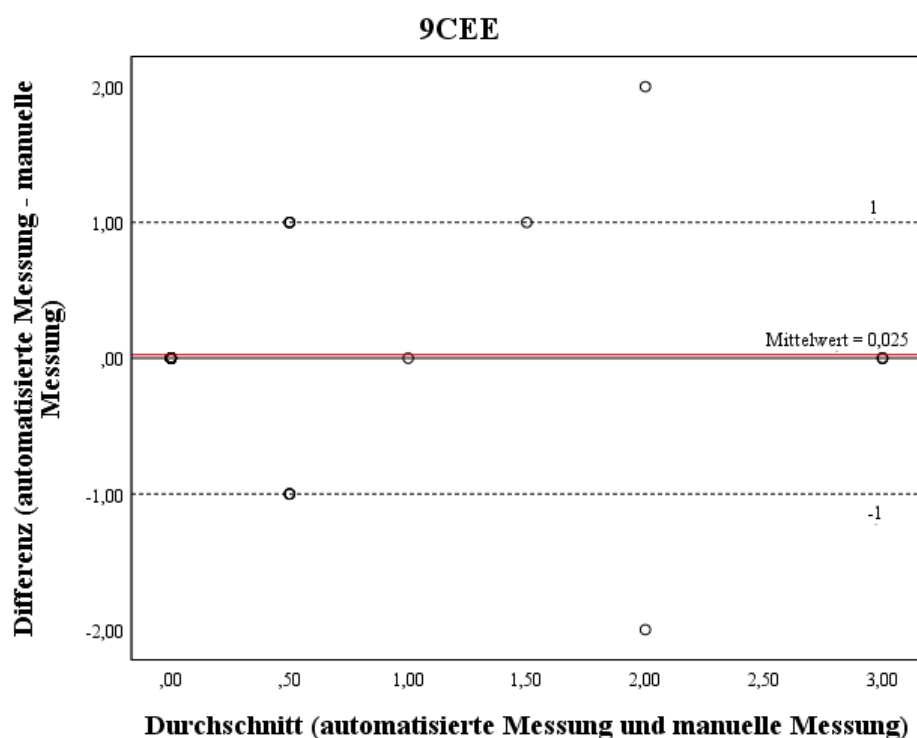


Abbildung 33: Bland-Altman-Diagramm der Messwerte für den 9cEE

Das Streudiagramm in Abbildung 32 zeigt bei allerdings nur geringer Zahl von 9cEEs eine starke Korrelation der Messergebnisse ($r = 0,786$, $p < 0,001$). Abbildung 33 gibt wieder, dass 95% der Messwertdifferenzen innerhalb eines Intervalls von ± 1 liegen (95%-Konfidenzintervall 73,6%-99,8%). Eine Abhängigkeit von Differenzen und Durchschnittswerten der Messpaare wurde nicht nachgewiesen ($r = 0,088$, $p = 0,645$). Eine weitere Interpretation der Daten erscheint bei nur seltenem Auftreten von 9cEEs nicht sinnvoll.

Die Datenverteilung für den 9cER für die manuellen und die automatischen Messungen zeigen Tabelle 16 und die Abbildungen 34 und 35.

Parameter	9cER	
	Manuelle Messung	Automatische Messung
Minimum	0%	0%
Maximum	1,75%	1,03%
Mittelwert	0,27	0,08
Standardabweichung	0,55	0,19
Variationskoeffizient	2,04	2,38

Tabelle 16: Statistische Kennwerte für den 9cER der manuellen und der automatischen Messungen

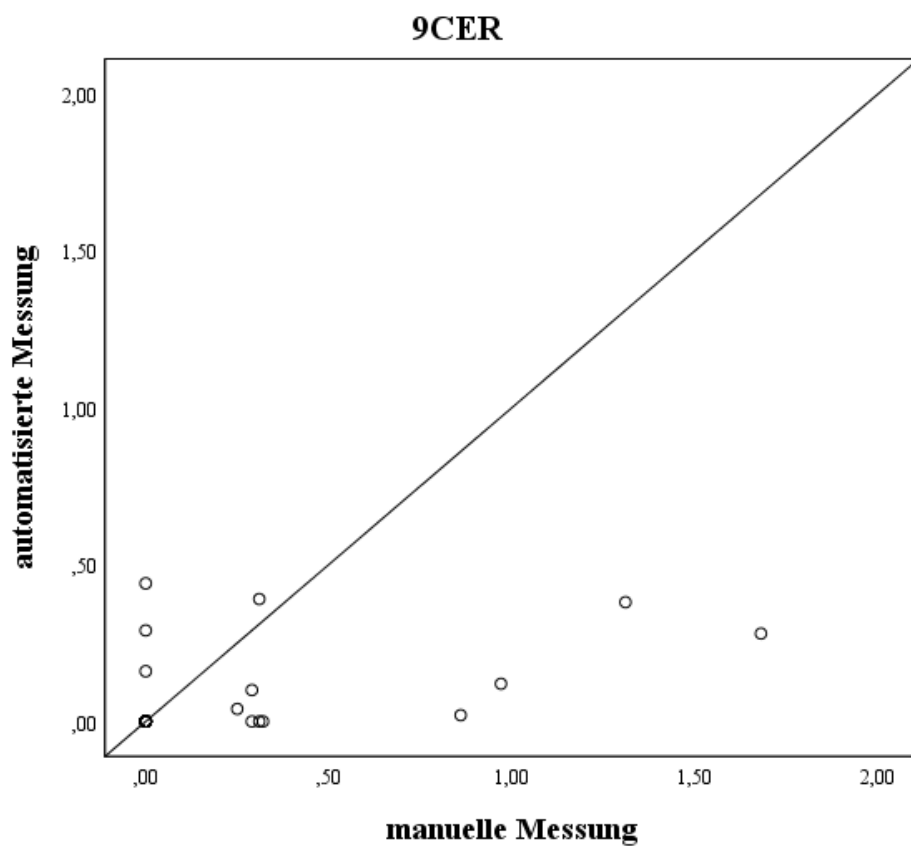


Abbildung 34: Streudiagramm der Messwerte für den 9cER

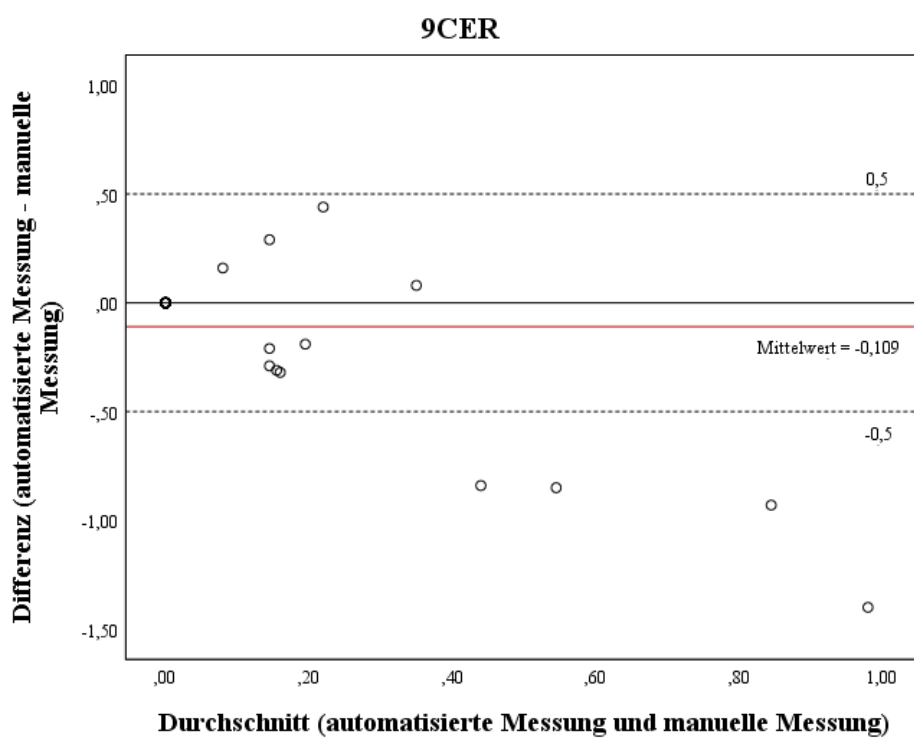


Abbildung 35: Bland-Altman-Diagramm der Messwerte für den 9cER

Das Streudiagramm in Abbildung 34 veranschaulicht, dass die 9cER-Werte für die automatische Messung in aller Regel klein sind und sich nur bis in den Bereich von 0,5% erstrecken, während sie für die manuelle Messung mit ihrer eingeschränkten Zahl von Analysezellen etwas höher liegen und bis in den 2%-Bereich reichen. Bei der Interpretation der Daten muss aber auch die Skalierung des Diagramms beachtet werden, das hier sehr hoch auflöst. Berücksichtigt man dieses, so erscheinen die Daten für den 9cER durchaus vom Trend her noch gut korreliert ($r = 0,494$, $p = 0,001$).

Das Bland-Altman-Diagramm (Abb. 35) weist aus, dass 90% der Differenzen von automatischen und manuellen Messungen beim 9cER innerhalb eines Intervalls von $\pm 0,5\%$ liegen. Das 95%-Konfidenzintervall beträgt hier 67,5%-97,2%. Es errechnet sich dabei eine Korrelation der Differenzen für den 9cER ($r_s = 0,449$, $p = 0,003$).

Betrachtet man die Ergebnisse von 9cEE und 9cER im Vergleich, so relativieren sich ähnlich wie beim Vergleich von 5cEE und 5cER die Unterschiede in den Daten bei der prozentualen Darstellung.

3.3.7 DNA-Stammlinie

Bei der manuellen Messung und bei der automatischen Messung konnten in allen 40 Fällen DNA-Stammlinien ermittelt werden, die bei der manuellen Messung zwischen 1,80c und 3,63c und bei der automatischen Messung zwischen 1,51c und 4,01c lagen. Die Ergebnisse beziehen sich bei euploiden Fällen jeweils auf die erste Stammlinie, bei aneuploiden Fällen auf die erste aneuploide Stammlinie. Weitere Stammlinien (bis zu drei), die im Histogramm zu sehen waren, stellten sich als Verdopplung der jeweils ersten Stammlinie dar.

Die entsprechende Datenübersicht ist in Tab. 17 wiedergegeben. In den Abb. 36 und 37 sind das zugehörige Streudiagramm und ein Bland-Altman-Diagramm dargestellt.

Parameter	DNA-Stammlinie	
	Manuelle Messung	Automatische Messung
Minimum	1,80c	1,51c
Maximum	3,63c	4,01c
Mittelwert	2,38c	2,16c
Standardabweichung	0,55c	0,52c
Variationskoeffizient	0,24	0,23

Tabelle 17: Statistische Kennwerte für die DNA-Stammlinie der manuellen und der automatischen Messungen

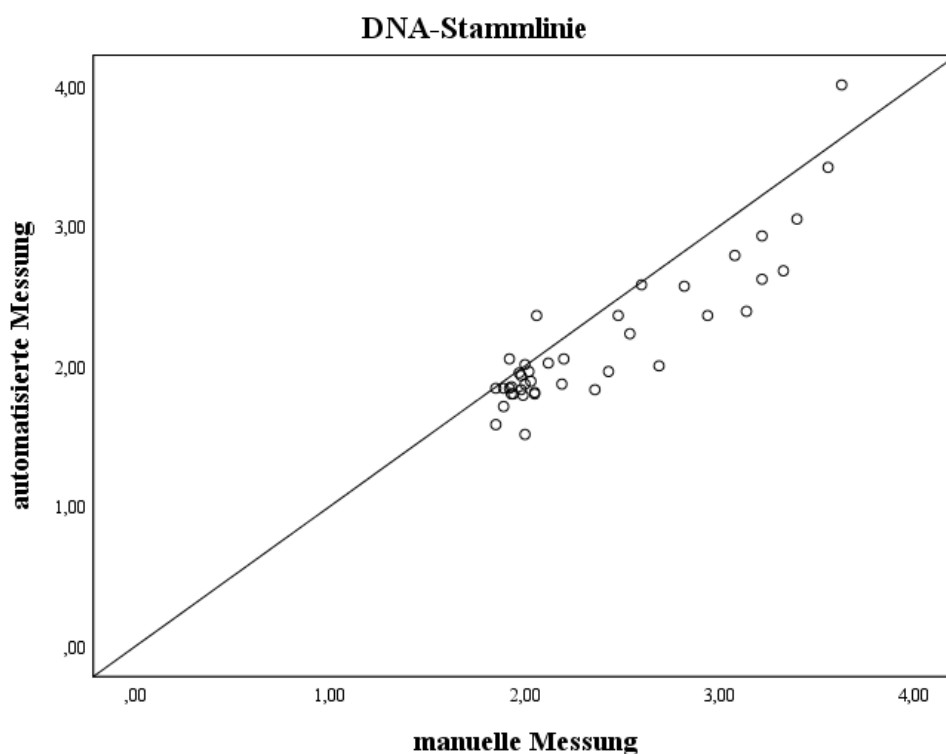


Abbildung 36: Streudiagramm der Messwerte der DNA-Stammlinie

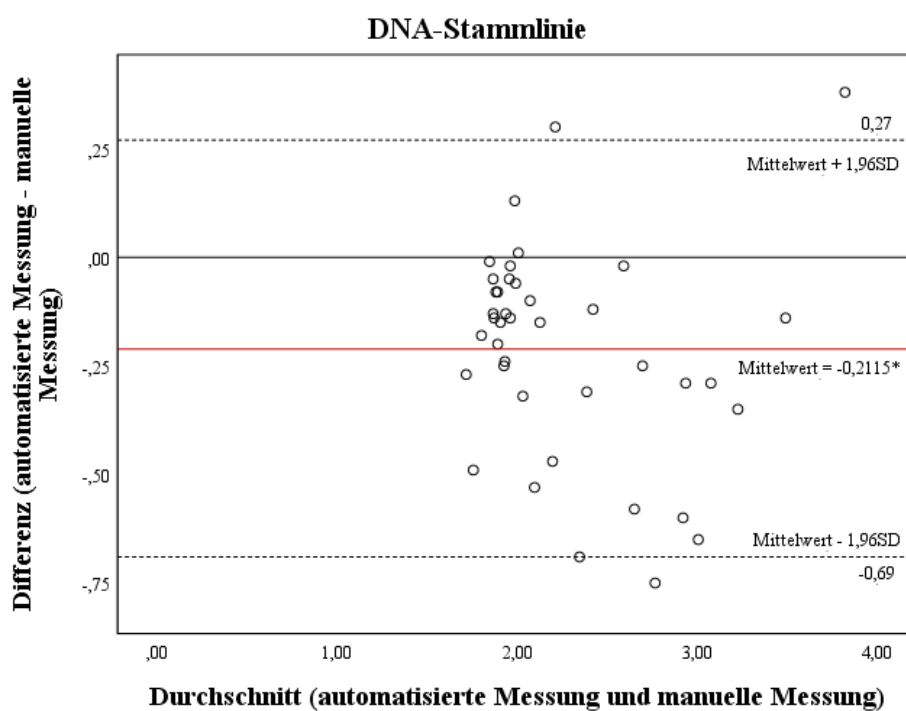


Abbildung 37: Bland-Altman-Diagramm der Messwerte der DNA-Stammlinien

Das Streudiagramm in Abb. 36 zeigt, dass die große Mehrzahl der Messwertpaare unter der Winkelhalbierenden liegt, so dass die manuellen Messungen etwas höhere Werte als die automatischen Messungen aufweisen. Im Mittel macht dieser Unterschied 0,22c aus. Dabei sind die Messergebnisse dennoch hoch korreliert ($r = 0,896$, $p < 0,001$), aber eben

nicht vollständig identisch. Dieser Effekt schlägt sich, wie unten in Kapitel 3.6 dargestellt und in Kapitel 4.3 diskutiert, auch in diagnostischen Diskrepanzen nieder. Eine Korrelation zwischen den Messwertdifferenzen und den durchschnittlichen Werten wurde nicht nachgewiesen ($r_s = -0,283$, $p = 0,077$).

3.4 Diagnostische Interpretation der manuellen und der automatischen DNA-Bildzytometrie

Im Folgenden werden die DNA-bildzytometrischen Messungen bezüglich der Frage ausgewertet, ob sie als DNA-aneuploid oder als DNA-euploid zu interpretieren sind. Die beiden Interpretationsmodi, nämlich die Auswertung nach der Lage der DNA-Stammlinie und nach der Zahl der 9cEE, wurden in Kapitel 2.9.2 vorgestellt und werden hier in den Kapitel 3.4.1 und 3.4.2 einzeln bzw. in Kapitel 3.4.3 kombiniert angewendet.

3.4.1 Auswertung nach Lage der DNA-Stammlinie

In Tabelle 18 sind die Messwerte für die DNA-Stammlinien (vgl. Kap. 3.3.7) diagnostisch ausgewertet.

Parameter	DNA-Stammlinie (automatische Messung)		
DNA-Stammlinie (manuelle Messung)	euploid	aneuploid	gesamt
euploid	20	5	25
aneuploid	3	12	15
gesamt	23	17	40

Tabelle 18: Übereinstimmungsmatrix für die DNA-Stammlinien-Auswertung bei manueller bzw. automatischer Messung, Chi²-Test: $p = 0,0002$

Eine Übereinstimmung in der Zuordnung einer euploiden DNA-Stammlinie lag in 20 Fällen (50%) vor, in der Zuordnung zu einer aneuploiden Stammlinie in 12 Fällen (30%). Acht Fälle (20%) wurden divergent beurteilt, wobei in 3 Fällen ausschließlich die manuelle Messung eine aneuploide Stammlinie aufwies. In 5 Fällen war die DNA-Stammlinie der manuellen Messung euploid, die der automatischen aneuploid. Der zugehörige Kappa-Wert betrug 0,584 (moderate agreement).

3.4.2 Auswertung nach Zahl der 9cEE

In Tabelle 19 sind die Messwerte für den 9cEE (vgl. Kap. 3.3.6) nach den Vorgaben der Einzelzellaneuploidie diagnostisch ausgewertet.

Parameter	9cEE (automatische Messung)		
9cEE (manuelle Messung)	euploid	aneuploid	gesamt
euploid	27	3	30
aneuploid	3	7	10
gesamt	30	10	40

Tabelle 19: Übereinstimmungsmatrix für die Auswertung des 9cEE nach den Vorgaben der Einzelzellaneuploidie bei manueller bzw. automatischer Messung, Chi²-Test: $p = 0,0001$

In 34 Fällen lag eine Übereinstimmung in der Zuordnung zu einer euploiden bzw. aneuploiden Zuordnung der 9cEE-Werte vor (85,0%). Sechs Fälle wurden divergent beurteilt, wobei in jeweils drei Fällen 9cEEs nur in einer der beiden Messungen vorkamen. Als Kappa-Wert wurde 0,60 ermittelt (moderate agreement).

3.4.3 Kombinierte Auswertung nach Lage der DNA-Stammlinie und Zahl der 9cEE

In Tabelle 20 sind die Messwerte in Kombination beider Variablen diagnostisch ausgewertet. Als aneuploid wurden Messungen angesehen, die eine aneuploide DNA-Stammlinie oder einen 9cEE > 0 aufwiesen.

Parameter	Kombinierte Auswertung (automatische Messung)		
Kombinierte Auswertung (manuelle Messung)	euploid	aneuploid	gesamt
euploid	12	6	18
aneuploid	5	17	22
gesamt	17	23	40

Tabelle 20: Übereinstimmungsmatrix für die kombinierte Auswertung von DNA-Aneuploidie bei manueller bzw. automatischer Messung, Chi²-Test: $p = 0,0052$

Eine Übereinstimmung in der Zuordnung zu einer euploiden bzw. einer aneuploiden DNA-Verteilung bei kombinierter Auswertung lag in 29 Fällen (72,5%) vor. Elf Fälle wurden divergent beurteilt, wobei in 5 Fällen die manuelle Messung im Gegensatz zur automatischen Messung aneuploid war; in 6 Fällen verhielt es sich anders herum. Der Kappa-Wert betrug 0,442 (moderate agreement).

Im Vergleich der drei Auswertemodi zeigt sich, dass die größte Übereinstimmung mit 85% in der Zuordnung von Aneuploidie und Euploidie beim Vergleich der 9cEE besteht. Die Lage der Stammlinien (80%) und die Kombination aus 9cEE und Lage der Stammlinien (72,5%) zeigten eine geringere Übereinstimmung.

3.5 Treffsicherheit der manuellen und der automatischen DNA-Bildzytometrie

Im folgenden Kapitel wird die Treffsicherheit der drei diagnostischen DNA-bildzytometrischen Parameter anhand der Korrelation mit den zytologischen Diagnosen und den histologischen Diagnosen dargestellt. Dazu werden zuerst die jeweiligen Diagnosen gruppiert dem Ergebnis der DNA-Bildzytometrie gegenübergestellt und danach mit den jeweiligen Messinterpretationen gemäß den Einteilungen in Kapitel 3.2 verglichen. Dabei soll von vornherein angemerkt werden, dass die geringe Fallzahl von 40 Fällen nicht zu verallgemeinernden Rückschlüssen aus den Ergebnissen dieses Teilaspektes unserer im Wesentlichen methodisch ausgerichteten Studie berechtigen kann.

3.5.1 Treffsicherheit nach Lage der DNA-Stammlinie

Für die Lage der DNA-Stammlinie und ihre Interpretation als euploid oder aneuploid zeigen die Tabellen 21 - 25 die entsprechenden Datenverteilungen.

Zytologische Diagnose

Zytologische Diagnose	DNA-Stammlinie			
	manuelle Messung		automatische Messung	
	euploid	aneuploid	euploid	aneuploid
negativ	5	2	6	1
zweifelhaft	17	8	16	9
dringender Verdacht	2	4	1	5
positiv	1	1	0	2
negativ	5	2	6	1
zw. / dV / pos.	20	13	17	16

Tabelle 21: Datenverteilung nach Lage der DNA-Stammlinie, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv

Zytologische Diagnose: "negativ" gegen "zw. / dV / pos."	DNA-Stammlinie	
	manuelle Messung	automatische Messung
Sensitivität	39,4%	48,5%
Spezifität	71,4%	85,7%
PPV	86,7%	94,1%
NPV	20,0%	26,1%
Gesamttreffsicherheit	45,0%	55,0%

Tabelle 22: Treffsicherheit nach Lage der DNA-Stammlinie, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv

Histologische Diagnose

Histologische Diagnose	DNA-Stammlinie			
	manuelle Messung		automatische Messung	
	euploid	aneuploid	euploid	aneuploid
Struma / reaktiv	15	0	13	2
Adenom	8	13	9	12
Karzinom	2	2	1	3
kein Tumor	15	0	13	2
Tumor	10	15	10	15
gutartig	23	13	22	14
bösartig	2	2	1	3

Tabelle 23: Datenverteilung nach Lage der DNA-Stammlinie, bezogen auf das Ergebnis der histologischen Untersuchung

Histologische Diagnose: "kein Tumor" gegen "Tumor"	DNA-Stammlinie	
	manuelle Messung	automatische Messung
Sensitivität	60,0%	60,0%
Spezifität	100%	86,7%
PPV	100%	88,2%
60,0%	60,0%	56,5%
Gesamttrefferlichkeit	75,0%	70,0%

Tabelle 24: Treffsicherheit nach Lage der DNA-Stammlinie, bezogen auf das Ergebnis der histologischen Untersuchung ("kein Tumor" gegen "Tumor")

Histologische Diagnose: "gutartig" gegen "bösartig"	DNA-Stammlinie	
	manuelle Messung	automatische Messung
Sensitivität	50,0%	75,0%
Spezifität	63,9%	61,1%
PPV	13,3%	17,7%
NPV	92,0%	95,7%
Gesamttrefferlichkeit	62,5%	62,5%

Tabelle 25: Treffsicherheit nach Lage der DNA-Stammlinie, bezogen auf das Ergebnis der histologischen Untersuchung ("gutartig" gegen "bösartig")

Fasst man die Ergebnisse zusammen, so zeigt sich im Vergleich beider Messmethoden ein zwar gut vergleichbares Bild, dass aber bezogen auf einige wichtige Daten etwas günstiger für die automatische Messung ausfällt.

Dies gilt sowohl für alle Vergleichswerte bei der Zytologie, aber auch bei der Interpretation der Messergebnisse im histologischen Vergleich zwischen "bösartig" und "gutartig". Beim Vergleich "Tumor" gegen "kein Tumor" schneidet hingegen die manuelle Messung bei den meisten Kenndaten etwas besser ab.

3.5.2 Treffsicherheit nach Zahl der 9cEE

Für die Zahl der 9cEE in der Interpretation als euploid oder aneuploid sind die Daten in den Tabellen 26 - 30 für Zytologie und Histologie wiedergegeben.

Zytologie

Zytologische Diagnose	9cEE			
	manuelle Messung		automatische Messung	
	euploid	aneuploid	euploid	aneuploid
negativ	7	0	5	2
zweifelhaft	18	7	19	6
dringender Verdacht	3	3	4	2
positiv	2	0	2	0
negativ	7	0	5	2
zw. / dV / pos.	23	10	25	8

Tabelle 26: Datenverteilung nach Zahl der 9cEE, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv

Zytologische Diagnose: "negativ" gegen "zw. / dV / pos."	9cEE	
	manuelle Messung	automatische Messung
Sensitivität	30,3%	24,2%
Spezifität	100%	71,4%
PPV	100%	80,0%
NPV	23,3%	16,7%
Gesamttreffsicherheit	42,5%	32,5%

Tabelle 27: Treffsicherheit nach Zahl der 9cEE, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv

Histologie

Histologische Diagnose	9cEE			
	manuelle Messung		automatische Messung	
	euploid	aneuploid	euploid	aneuploid
Struma / reaktiv	12	3	12	3
Adenom	15	6	15	6
Karzinom	3	1	3	1
kein Tumor	12	3	12	3
Tumor	18	7	18	7
gutartig	27	9	27	9
bösartig	3	1	3	1

Tabelle 28: Datenverteilung nach Zahl der 9cEE, bezogen auf das Ergebnis der histologischen Untersuchung

Histologische Diagnose: "kein Tumor" gegen "Tumor"	9cEE	
	manuelle Messung	automatische Messung
Sensitivität	28,0%	28,0%
Spezifität	80,0%	80,0%
PPV	70,0%	70,0%
NPV	40,0%	40,0%
Gesamttrefferlichkeit	47,5%	47,5%

Tabelle 29: Treffsicherheit nach Zahl der 9cEE, bezogen auf das Ergebnis der histologischen Untersuchung ("kein Tumor" gegen "Tumor")

Histologische Diagnose: "gutartig" gegen "bösartig"	9cEE	
	manuelle Messung	automatische Messung
Sensitivität	25,0%	25,0%
Spezifität	75,0%	75,0%
PPV	10,0%	10,0%
NPV	90,0%	90,0%
Gesamttrefferlichkeit	70,0%	70,0%

Tabelle 30: Treffsicherheit nach Zahl der 9cEE, bezogen auf das Ergebnis der histologischen Untersuchung ("gutartig" gegen "bösartig")

Die Ergebnisse zeigen im Vergleich beider Messmethoden bei der Zytologie etwas bessere Kennwerte für die manuelle Messung, bei der Histologie sind die Werte insgesamt zwischen beiden Messreihen identisch, obwohl es innerhalb von Fällen (siehe Kapitel 3.6) durchaus diskrepante Ergebnisse für den 9cEE in der manuellen und der automatischen Messung gab; zahlenmäßig gleichen sich die Diskrepanzen aber genau aus.

3.5.3 Treffsicherheit nach Kombination aus der Lage der DNA-Stammlinie und der Zahl der 9cEE

Für die Kombination aus Lage der DNA-Stammlinie und Zahl der 9cEE und ihre diagnostische Interpretation sind die Daten und Auswertungen in den Tabellen 31 - 35 gezeigt.

Zytologie

Zytologische Diagnose	Kombination aus DNA-Stammlinie und 9cEE			
	manuelle Messung		automatische Messung	
	euploid	aneuploid	euploid	aneuploid
negativ	5	2	4	3
zweifelhaft	11	14	13	12
dringender Verdacht	1	5	0	6
positiv	1	1	0	2
negativ	5	2	4	3
zw. / dV / pos.	13	20	13	20

Tabelle 31: Datenverteilung der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der zytologischen Untersuchung

Zytologische Diagnose: "negativ" gegen "zw. / dV / pos."	Kombination aus DNA-Stammlinie und 9cEE	
	manuelle Messung	automatische Messung
Sensitivität	60,6%	60,6%
Spezifität	71,4%	57,1%
PPV	90,9%	87,0%
NPV	27,8%	23,5%
Gesamttreffsicherheit	62,5%	60,0%

Tabelle 32: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv

Histologie

Histologische Diagnose	Kombination aus DNA-Stammlinie und 9cEE			
	manuelle Messung		automatische Messung	
	euploid	aneuploid	euploid	aneuploid
Struma / reaktiv	12	3	10	5
Adenom	5	16	7	14
Karzinom	1	3	0	4
kein Tumor	12	3	10	5
Tumor	6	19	7	18
gutartig	17	19	17	19
bösartig	1	3	0	4

Tabelle 33: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der histologischen Untersuchung

Histologische Diagnose: "kein Tumor" gegen "Tumor"	Kombination aus DNA-Stammlinie und 9cEE	
	manuelle Messung	automatische Messung
Sensitivität	76,0%	72,0%
Spezifität	80,0%	66,7%
PPV	86,4%	78,2%
NPV	66,7%	58,8%
Gesamttreffsicherheit	77,5%	70,0%

Tabelle 34: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der histologischen Untersuchung ("kein Tumor" gegen "Tumor")

Histologische Diagnose: "gutartig" gegen "bösartig"	Kombination aus DNA-Stammlinie und 9cEE	
	manuelle Messung	automatische Messung
Sensitivität	75,0%	100%
Spezifität	47,2%	47,2%
PPV	13,6%	17,4%
NPV	94,4%	100%
Gesamttreffsicherheit	50%	52,5%

Tabelle 35: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der histologischen Untersuchung ("gutartig" gegen "bösartig").

Die Ergebnisse weisen im Vergleich beider Messmethoden bei der Zytologie etwas bessere Kennwerte für die manuelle Messung auf; Gleiches gilt bei der Histologie für den Vergleich "Tumor" gegen "kein Tumor". Bei der Differenzierung zwischen gut- und bösartig schneidet die automatische Messung minimal besser ab.

3.6 Diskrepante Fälle

Im Folgenden sollen die in der diagnostischen Auswertung diskrepanten Fällen vorgestellt werden. In der Diskussion wird dann in Kapitel 4.4 versucht, sie systematisch begründet in Gruppen einzuteilen.

Dabei ist aus methodischer und inhaltlicher Sicht nicht wichtig, dass es Fälle geben kann, die mit beiden Messmethoden übereinstimmend z.B. nur über die DNA-Stammlinie als aneuploid erkannt werden, nicht hingegen über den 9cEE - oder umgekehrt. Wesentlich sind hingegen die Fälle, in denen sich bezogen auf den gleichen Parameter die Ergebnisse der manuellen und der automatischen Messung unterscheiden. Sie sind in der folgenden Tabelle 36 wiedergegeben. Dabei zeigten sich Unterschiede in der Interpretation der DNA-Stammlinie in acht Fällen (20%) und für den 9cEE in sechs Fällen (15%). Für die kombinierte Auswertung ergaben sich Unterschiede in elf Fällen (27,5%).

Patienten-Nr.	DNA-Stammlinie manuell	DNA-Stammlinie automatisch	9cEE manuell	9cEE automatisch
3	2,02c	1,96c	0	1
4	3,33c	2,68c	1	0
7	2,06c	2,36c	0	0
9	3,63c	4,01c	0	1
10	2,69c	2,00c	0	0
14	2,00c	1,51c	4	6
17	1,85c	1,58c	0	0
23	2,36c	1,83c	0	0
25	1,89c	1,71c	0	0
28	1,99c	1,79c	0	0
31	1,85c	1,84c	1	0
34	2,48c	2,36c	0	1
36	1,93c	1,85c	1	0
38	2,43c	1,96c	0	0

Tabelle 36: Datenverteilung der DNA-Stammlinien und der 9cEE-Werte für die manuelle und die automatische Messung bei den 14 interpretatorisch diskrepanten Fällen. Werte, die die Kriterien für DNA-Aneuploidie erfüllen, sind fettgedruckt wiedergegeben. Im Endergebnis der kombinierten Auswertung diskrepante Fälle sind rot dargestellt.

4 Diskussion

Die Feinnadelaspirationsbiopsie (FNAB) der Schilddrüse wird in Verbindung mit der Zytopathologie seit langem als Referenzmethode zur Beurteilung fokaler Läsionen der Schilddrüse eingesetzt.

Die in der Literatur angegebenen Daten zur Sensitivität von 65-98% und zur Spezifität von 72-100% (Haider et al., 2011), und die Raten von falsch negativen (0,7 –16,1%) (Goldstein et al., 2002) und falsch positiven Diagnosen (6 und 11%) (Lewis et al., 2009), (Wang et al., 2011) spiegeln dabei auch das Problem wider, das sich aus der eingeschränkten Zuordbarkeit zellreicher follikulärer Punktate ergibt. Je nachdem nämlich, ob man diese Zellbilder, die in ca. 5% bis 20% der Punktate auftreten (Pasha et al., 2020), (Seagrove-Guffey et al., 2018) bei der Dichotomisierung als "negativ" oder "positiv" einordnet, ergeben sich entsprechende Konsequenzen für die Treffsicherheit. Jedenfalls ist es in der konventionellen Schilddrüsenzytologie aufgrund dieser diagnostischen Einordnungsprobleme nicht leicht möglich, sowohl eine hohe Sensitivität, als auch eine hohe Spezifität zu erreichen. Gleiches gilt sinngemäß auch für PPV und NPV.

In der klinischen Anwendung soll es bei der Schilddrüsenzytologie vor allem darum gehen, die Zahl falsch-negativer Fälle gering zu halten. Denn wenn das Ziel ist, über eine negative Zytologie nicht notwendige operative Eingriffe an gutartigen Knoten zu vermeiden, dann muss eine hohe Spezifität im Vordergrund stehen (Paschke et al., 2011). Die Zahlen in Publikationen weisen dabei das Problem auf, dass der wirkliche Anteil falsch negativer Diagnosen kaum bestimmt werden kann, da nur in ca. 10% der zytologisch benignen Fälle eine operative Klärung herbeigeführt wird und somit eine Histologie zum Vergleich vorliegt. Der Preis für die angestrebte möglichst geringe Zahl falsch negativer Fälle ist dann die Inkaufnahme einer etwas vermehrten Rate falsch Positiver (Bakhos et al., 2000). Ziel der zytologischen Diagnostik ist also primär, eine für das weitere klinische Vorgehen klare Entscheidung zu treffen und die Zahl der „dringend verdächtigen“ Präparate möglichst gering zu halten, ohne damit die Sensitivität des Tests zu verringern (Renshaw, 2011).

Bezüglich dessen, was in der Schilddrüsenzytologie aktuell möglich ist, kann zur Veranschaulichung auf Zahlen einer 2020 publizierten Meta-Analyse verwiesen werden (Lan et al., 2020), die sich auf gepoolte Ergebnisse von 4927 Schilddrüsenzytologien aus 10 Studien bezieht. Hiervon waren 3031 richtig negativ, 1409 richtig positiv, 451 falsch negativ und 36 falsch positiv. Hieraus errechnen sich Werte für die Sensitivität von 75,8%, für die Spezifität von 98,9%, für den PPV von 97,5% und für den NPV von 87,0%. In der Arbeit wurden ferner als gepoolte Sensitivität 72% und als gepoolte Spezifität 99% angegeben.

Insgesamt belegen die Daten, dass Methoden- und Organ-immanent eine Treffsicherheit der konventionellen Schilddrüsenzytologie in den sonst üblichen Dimensionen der Zytologie wohl nicht möglich ist, solange man dem diagnostischen Problem der "follikulären Läsionen" nicht Herr wird. Die Frage, die sich hieran anschließend stellt, ist, inwieweit adjuvante morphologische oder molekulare Methoden sinnbringend angewandt werden könnten, um zu präziseren Aussagen zu gelangen.

Die vorliegende Arbeit geht dabei der Frage nach, inwieweit die DNA-Bildzytometrie als adjuvante Methode eingesetzt werden könnte, wobei das Hauptaugenmerk auf dem methodischen Vergleich zwischen der etablierten manuellen DNA-Bildzytometrie und einem neuen automatischen Verfahren liegt. Zum Vergleich wurden sowohl unterschiedliche Parameter gegenübergestellt als auch die allgemeine Treffsicherheit beider Verfahren ermittelt.

Zunächst soll aber allgemein auf die Möglichkeiten und Grenzen der Anwendung adjuvanter Methoden in der Schilddrüsenzytologie eingegangen werden und die DNA-Bildzytometrie damit als eine unter mehreren methodischen Ansätzen eingeordnet werden.

4.1 Adjuvante Methoden in der Schilddrüsenzytologie

Adjuvante Methoden haben sich nicht nur in der Histopathologie, sondern auch in der diagnostischen Zytopathologie inzwischen weithin etabliert. Damit sind allgemein Sonderverfahren gemeint, die nicht routinemäßig bei jedem Fall appliziert werden, sondern ergänzend hinzugenommen werden, wenn im Rahmen der Befundung ein Problem gesehen wird, das man nicht mit der konventionellen Diagnostik lösen kann. Es bedarf also einer speziellen fallbezogenen Indikation.

Zu den adjuvanten Methoden gehören neben konventionellen Sonderfärbungen, auf die ebenso wie auf die diagnostisch kaum noch eingesetzte Elektronenmikroskopie nicht eingegangen werden soll, insbesondere die Immunzytochemie, molekularpathologische Verfahren und die DNA-Bildzytometrie.

Immunzytochemie

Am wichtigsten und auch am weitesten in der Pathologie verbreitet ist dabei die Immunzytochemie an Ausstrich- oder Zentrifugationspräparaten bzw. an aus sogenannten Zellblöcken gewonnenen histologischen Schnittpräparaten. Hierbei werden zytologische Materialien in einen Zellblock verbracht und dann wie eine normale Histologie weiterverarbeitet.

An Schilddrüsenpunktaten werden immunzytochemische Verfahren allerdings nur relativ selten vorgenommen. Dabei geht es zum einen um die Identifikation der zu untersuchenden Zellen als Schilddrüsen-eigen. Gängige Marker wären hier ins-

besondere TTF-1 zur Identifikation von ortsständigen Epithelien der Schilddrüse, ferner Thyreoglobulin und Calcitonin zur Differenzierung zwischen Thyreozyten und C-Zellen. Außerdem können als Marker Parathormon oder endokrine Marker wie Synaptophysin sinnvoll sein, mit denen sich Nebenschilddrüsenepithelien abgrenzen lassen, welche Thyreozyten sehr ähnlich sehen können und Schilddrüsenpunktaten beigemischt sein können.

Zum anderen wurden Konzepte entwickelt, um follikuläre Läsionen der Schilddrüse als benigne bzw. maligne zu identifizieren. Als immunhistochemische Marker zur Abklärung solcher unklarer Schilddrüsenläsionen werden in der Histologie und in der Zytologie seit Jahren insbesondere die Marker Galektin 3, CK19, Thyreoperoxidase (TPO) und HBME-1 eingesetzt, mit denen sich durchaus richtungsweisende, nicht aber für Malignität oder eine bestimmte Entität beweisende Erkenntnisse gewinnen lassen:

- So wurde 2008 bereits über die Wertigkeit von Galektin-3 an Schilddrüsenpunktaten berichtet und ermittelt, dass 280 von 331 für Galektin-3 negative Punktate eine gutartige Histologie hatten (84,6%), während 101 der 134 Galektin-3-positiven Fälle sich als maligne erwiesen (75,4%) (Bartolazzi et al., 2008).
- In einer anderen Studie aus dem Jahr 2012 wurden Galektin-3, CK19, TPO und HBME-1 eingesetzt (Paunovic et al., 2012); mit allen vier Markern betrug die Sensitivität zur Identifikation von Schilddrüsenkarzinomen zwischen 66,5% und 82,2%. Insbesondere eine kombinierte Anwendung von HBME-1 und TPO wurde als sinnvoll in der Differentialdiagnostik angesehen.
- An der Markerauswahl und der immer noch bestehenden Relativität der Aussagen hat sich bis heute nichts geändert. Auch eine ganz aktuelle Arbeit aus dem Jahr 2021 an 224 Zellblöcken verwendete nach wie vor die Marker CK19, HBME-1 und Galektin-3 und ermittelte Treffsicherheiten zur Detektion eines malignen Tumors von 65,1%, 72,1% bzw. 70,3% (Abram et al., 2021).

Immunzytochemisch ist somit eine gewisse Wahrscheinlichkeitsaussage möglich, so dass man beispielsweise eine zytologische Verdachtsdiagnose auf einen Tumor bei entsprechender Markerkonstellation verdichten kann. Beweischarakter kommt der Immunzytochemie allerdings nicht zu.

Molekularpathologie

Zytologische Präparate sind gut für die Anwendung PCR-basierter oder FISH-basierter molekularpathologischer Fragestellungen geeignet. Sowohl luftgetrocknete Präparate als auch Alkohol-fixiertes Material verfügt üblicherweise über einen guten Erhalt von DNA, oft besser als nach Formalin-Fixation für die Histologie (Grote et al. 2003).

Wesentliche genetische Veränderungen an papillären Schilddrüsenkarzinomen beziehen sich auf die BRAF V600E Mutation und Fusionen im RET-Gen, während Mutationen im RAS-Gen und weiteren RAS-assoziierten Genen (BRAF K601E Mutation, PTEN-Mutationen etc.) bei follikulären Neoplasien typisch sind. Inzwischen werden derartige Untersuchungen auch in Multigentests angewandt. Ihr Stellenwert wurde 2020 in einem aktuellen Review bezogen auf drei gewerblich erhältliche Multigentests (ThyroSeqv3, ThyGenX/ThyraMIR und Afirma GSC) zusammengefasst (Nishino & Krane, 2020):

- In den Validierungsstudien wiesen die drei Panels Werte für die Sensitivität von 89%-94%, für die Spezifität von 68%-82%, für den PPV von 47%-74% und für den NPV von 94%-97% auf, angewandt an zytologischen Präparaten mit dem Bild einer unklaren oder tumorverdächtigen follikulären Läsion.
- In postoperativen Postvalidierungsstudien ergab sich für den Multigentest ThyroSeqv2, eine Vorgängerversion des ThyroSeqv3, eine Sensitivität von 73,2%, eine Spezifität von 70,9%, ein PPV von 69,4% und ein NPV von 93,8%. Für den Multigentest Afirma GSC betrugen die vier Werte 97,4%, 17,8%, 27,9% und 95,4%.

Diese Werte liegen somit überwiegend höher als die der konventionellen Zytologie und geben somit durchaus Anhaltspunkte für die Dignität der zugrunde liegenden Veränderungen, können aber das Problemfeld der schwierigen Dignitätsbestimmung follikulärer Läsionen ebenfalls noch nicht lösen. Ähnliche Ergebnisse zeigen auch andere Studien, in denen kleinere Genpanel verwendet worden sind (Labourier et al., 2015) (Moore et al., 2017).

DNA-Bildzytometrie

Die DNA-Bildzytometrie ist eine effektive Methode in der Abgrenzung zwischen maligne transformierten und benignen Läsionen und wird insbesondere auch in der Cytopathologie in Düsseldorf seit langem betrieben und diagnostisch weiterentwickelt.

An der Schilddrüse liegen aus der Literatur allerdings nur relativ wenige Erfahrungen zur DNA-Bildzytometrie vor. Sie zeigen allgemein, dass auch mit dieser Methode, die bei anderen Organen sehr gut zwischen "Gut" und "Böse" diskriminieren kann, keine 100%-ige Lösung der Probleme in der Differentialdiagnostik follikulärer Läsionen erzielt werden kann.

Fasst man die Ergebnisse einer größeren, grundlegenden Studie aus der Cytopathologie in Düsseldorf aus der wissenschaftlichen Arbeitsgruppe von Böcking (Nadjari et al. 1999) an 162 Fällen von Schilddrüsenpunktaten im Kontext mit dem übrigen Wissen um das biologische Verhalten von Schilddrüsenepithelien in der Zellreplikation zusammen (zusammengefasst bei Biesterfeld et al. 1994), so ergibt sich in etwa folgendes Bild:

- Schilddrüsenepithelien unterliegen dem Phänomen der euploiden Polyploidisierung (Biesterfeld et al. 1994), können also DNA-Gehalte aufweisen, die in ruhenden Zellen Vielfachen von 2c entsprechen (4c, 8c, 16c ...), wobei der bisher historisch höchste publizierte Wert eines Schilddrüsenepithels von 128c ebenfalls in der Cytopathologie in Düsseldorf, aber noch in der Zeit der wissenschaftlichen Arbeitsgruppe von Pfitzer beschrieben wurde (Heide, 1982). Solche Werte entsprechen dem DNA-Gehalt von Zellen in der G₂/M'-, G₂/M"-, G₂/M'''-Phase usw.; sie sind insgesamt selten und werden von Phase zu Phase in geringerer Zahl vorkommen. Noch seltener sind dann Werte in den zwischen den verschiedenen G₂/M-Phasen liegenden Proliferationsphasen S', S'', S''' etc.

Ein DNA-Messwert oberhalb von 5c oder 9c kann somit nicht *à priori* als pathologisch aufgefasst werden. Dennoch wird nicht selten, auch an anderen polyploidisierenden Geweben wie z. B. dem Mesothel, das Auftreten von 9cEEs als Kriterium für Einzelzellaneuploidie herangezogen, gerade weil es sich bei seinem Auftreten um sogenannte *rare events* handelt, die von Stufe zu Stufe der Polyploidisierung immer seltener werden.

- Nicht-tumoröse Veränderungen, also alle Formen von Strumen, egal ob hyperthyreot, euthyreot oder hypothyreot, und auch entzündliche Läsionen wiesen bildzytometrische keine DNA-Stammlinienaneuploidie auf. In den insgesamt 89 gutartigen und bösartigen Schilddrüsentumoren waren aneuploide DNA-Stammlinien verbreitet und insgesamt etwa gleich häufig; sie fanden sich in 55% der Schilddrüsenadenome und in knapp 60% der malignen Tumoren (Nadjari et al. 1999). Es wurde aus den Ergebnissen geschlossen, dass DNA-Aneuploidie an der Schilddrüse nicht wie an anderen Organen als Marker für maligne Transformation oder Malignität, aber als Marker für Neoplasie dienen kann.

Auch durchflusszytometrisch kam man zu vergleichbaren Ergebnissen bei der Bestimmung des DNA-Index (Äquivalent zur DNA-Stammlinie) von Schilddrüsentumoren. Von 60 follikulären Adenomen waren 75% diploid, 13% tetraploid oder polyploid und 12% aneuploid, bei den 64 follikulären Karzinomen lagen die Werte bei 45%, 25% und 30% (Grant et al., 1990). Es ergaben sich damit zwar graduelle Unterschiede zur Nadjari-Studie, aber letztendlich konnte auch hier eine diagnostische Aussagekraft aus dem Nachweis von DNA-Aneuploidie im Hinblick auf die Dignitätsbestimmung eines Schilddrüsentumors nicht abgeleitet werden.

Die Idee, dass gutartige Läsionen auch an der Schilddrüse nicht aneuploid sind, wurde nur in wenigen Arbeiten mit untersucht und dort nicht durchgehend bestätigt. So waren in zwei kleineren durchflusszytometrischen Arbeiten zwar alle zehn Fälle (Kashyap et al., 1992) bzw. 17 Fälle (Horii et al., 1999) von mit untersuchten Strumen DNA-diploid. In einer größeren durchflusszytometrischen Studie an 224 Fällen von Thyreoidektomien

hingegen wiesen 50 Strumen DNA-Aneuploidie auf (22,3%) (Francia et al., 1997). Die genauere Durchsicht dieser Arbeit zeigt aber, dass methodische Vorbehalte insbesondere zur Eichung und "Einstellung" des diploiden Peaks angebracht sind, so dass der wirkliche Wert dieser Arbeit fraglich ist - insbesondere ist zu fragen, weswegen sich in größerer Zahl unerwartet hypodiploide "aneuploide" Stammlinie nachweisen ließen. Ähnliche methodische Einschränkungen aus alter Zeit gelten für eine durchflusszytometrische Arbeit an 74 Strumen und 50 Fällen von Morbus Basedow bzw. Hashimoto-Thyreoiditis, von denen ca. 20% als DNA-aneuploid eingeordnet wurden; auch wenn einige dieser Einstufungen zweifelhaft erscheinen, bleiben andere wiederum durchaus glaubhaft (Salmon et al., 1993). Andererseits beschreibt auch eine Arbeit aus der etablierten DNA-zytometrischen Gruppe um Sobrinho-Simões DNA-Aneuploidie in Strumen in einer Häufigkeit von 18,5% (15/81), allerdings an histologischen Schnitten durchgeführt (Castro et al., 2001), so dass auch hier methodische Einschränkungen aufgrund der Verwendung eines heute als obsolet angesehenen Verfahrens berücksichtigt werden müssen. Insgesamt haben die Resultate von Nadjari weiter Bestand, sollten aber möglicherweise etwas relativiert gesehen werden und nicht ohne weiteres verallgemeinert werden.

4.2 Methodischer und diagnostischer Vergleich der manuellen und der automatischen DNA-Bildzytometrie

Die Automatisierung von Prozessen hat auch vor der Pathologie in den letzten Jahren nicht Halt gemacht, da nur auf dem Boden einer möglichst hohen Standardisierung inhaltlich gute Ergebnisse und auf der Basis eines hohen Probendurchsatzes Untersuchungszahlen erreicht werden können, die den Einsatz einer Methode auch im diagnostischen Routinebetrieb erlauben.

Dies gilt gleichermaßen für molekularpathologische, immunhistochemische und bildanalytische Anwendungen, ferner auch für die Routinefärbungen in der Pathologie. Alle diese Methoden wurden früher in Handarbeit umgesetzt, während sie heute in Prozessierungs- oder Färbeautomaten bzw. kleinen, aber sehr effektiven PCR-Cyclern ablaufen.

Bei der DNA-Bildzytometrie ist seit vielen Jahren der Ablauf der DNA-spezifischen Feulgen-Färbung bereits auf Färbeautomaten umgestellt, was nach allgemeiner Meinung die Färbequalität verbessert hat, möglicherweise weil bei der Inkubation der Temperatur-abhängigen Färbeschritt weniger Variabilität herrscht. Die eigentliche Messung wird aber nach wie vor manuell durchgeführt, indem der Untersucher die Zellkerne, die gemessen werden sollen, mit einer Maus auf dem Bildschirm anklickt und ihre Zuordnung zu Referenz- oder Analysezellen ebenfalls manuell vorwählt. Dieses Verfahren ist gut und bewährt und diagnostisch auswertbar; in Düsseldorf werden pro Jahr mit der konventionell ausgeführten DNA-Bildzytometrie zwischen 500 und 800 Fälle in der Krankenversorgung bearbeitet.

Die hier vorliegende Studie befasst sich nun mit der Leistungsfähigkeit eines Prototyps eines automatischen DNA-bildzytometrischen Messsystems, das in einer Kooperation der Cytopathologie am UKD (Böcking, Leiter 1994-2009) und dem Lehrstuhl für Bildverarbeitung der RWTH Aachen (Aach, Meyer-Ebrecht) sowie einem Industriepartner aus China (Fa. Motic, Xiamen) entwickelt wurde. Ziel war, durch eine komplette Erfassung aller Zellen auf einem Präparat eine hohe Repräsentativität der Zellbilder zu erhalten und durch die vorherige Programmierung eines *Classifier*-Systems zu ermöglichen, dass das Messsystem selber die erfassten Zellkerne zu bestimmten Zellsorten (Granulozyten, Lymphozyten, Makrophagen, normale Epithelien, abnorme Epithelien, Artefakte) zuordnen kann. So bliebe für den Untersucher dann "nur" noch die Aufgabe, eine Überprüfung der Messung anhand einer Bildgalerie, in der die Bilder der gemessenen Zellkerne abgelegt werden, vorzunehmen und den Befund auszufertigen. Grundlegendes zur Bildanalyse und zur Generierung der *Classifier* wurde bereits in einer naturwissenschaftlichen Dissertation an der RWTH Aachen veröffentlicht (Friedrich, 2016).

Dabei ist es nicht möglich, einen allgemeingültigen *Classifier* zu entwickeln, mit dem sich alle Arten von Präparaten angehen ließen. Das hat zum einen damit zu tun, dass sich auffällige Zellveränderungen von Organ zu Organ etwas unterscheiden, zum anderen damit, dass auch die Präparation der Proben in der Cytopathologie organabhängig erfolgt. Die Art der Präparation wiederum hat dabei großen Einfluss auf Größe, Form und Struktur der Zellkerne, was dann vom jeweiligen *Classifier* berücksichtigt werden muss.

Bei der Erprobung der automatischen DNA-Bildzytometrie und dem Vergleich mit der manuellen DNA-Bildzytometrie handelt es sich hier um die dritte Studie.

Bei der ersten Arbeit wurden Messungen an Prostatakarzinomen vorgenommen (Methodik: Herstellung zytologischer Ausstrichpräparate nach enzymatischer Zellvereinzelung aus Paraffinblöcken) (Rezaie, 2018), bei der zweiten Arbeit Messungen an Körperhöhlenergüssen (Methodik: Ungefärbte luftgetrocknete Sedimentausstriche von unfixierten Ergusspunktaten) (Feider, 2018), (Böcking et al., 2019). Auf Gemeinsamkeiten und Unterschiede in den Ergebnissen wird in Kapitel 4.4 eingegangen.

Die hier vorliegende Studie geht wiederum von einem anderen Ansatz aus, nämlich von zuvor gefärbten, bereits als fertige Objektträger eingesandten luftgetrockneten Ausstrichen von Schilddrüsenpunktaten. Da es für den Prototypen des automatischen Messsystems keinen eigenen *Classifier* für Schilddrüse gibt und die Weiterentwicklung der Automatisierung an einem anderen Messsystem erfolgt, das in Düsseldorf nicht zur Verfügung steht, wurde für die vorliegende Untersuchung der *Classifier* für Mesothel, der auch an der Ergussarbeit angewandt wurde, verwendet. Dies erschien gerechtfertigt, da sich die zytologischen Besonderheiten der Schilddrüsenzytologie vielfach ähnlich auch beim Mesothel darstellen.

4.2.1 Vergleichbarkeit der Kenngrößen der manuellen und der automatischen DNA-Bildzytometrie

Insgesamt wurden in unserer Studie an 40 Messpaaren jeweils acht Grundgrößen und zwei abhängige Größen (in Klammern) ermittelt, nämlich die Zahl der Analysezellen, die Zahl der Referenzzellen, die mittlere Referenzzell-IOD, der minimale und der maximale C-Wert, der 5cEE (5cER), der 9cEE (9cER) und die Lage der DNA-Stammlinie. Im Folgenden sollen wesentliche Aspekte zu Übereinstimmungen und Unterschieden herausgearbeitet werden.

Dass sich bezüglich der Zahl der Analyse- und der Referenzzellen klare Unterschiede zwischen den Messpaaren ergeben haben, ist nicht überraschend und im Grunde Natur der Sache. Die Mittelwerte lagen in der automatischen Messreihe um den Faktor 7,9 bzw. 10,2 höher, was zeigt, dass die manuellen Messungen eben doch nur einen relativ kleinen Anteil der Zellen auf dem Objektträger abbilden. Das muss aber solange kein Nachteil sein, wie die Zellen der zugrunde liegenden Läsion in repräsentativem Maße erfasst sind. Dafür, dass dieses so ist, spricht die Erfahrung, dass sich im klinischen Verlauf die Ergebnisse manueller DNA-zytometrische Messungen fast nie als falsch positiv oder falsch negativ herausgestellt haben.

Die Werte für die Referenzzell-IOD unterschieden sich erstaunlicherweise nur gering. Der Mittelwert der automatischen Messungen (307,27) lag bei 94,8% des Mittelwertes der manuellen Messungen (324,27); entsprechend ist die Ausbleichung der Präparate jedenfalls im Mittel nur als gering anzusehen. Dies bedeutet, dass eine jahrelange Lagerung bereits DNA-bildzytometrisch ausgewerteter Präparate in geschlossenen Mappen, also in Dunkelheit, durchaus einen Erhalt der Feulgen-Färbung gewährleisten konnte. Allerdings muss in diesem Zusammenhang berücksichtigt werden, dass zwei Fälle, die in die Studie aufgenommen werden sollten, nicht mehr geeignet waren, eben weil wegen zu starker Ausbleichung die Zellkerne vom automatischen System nicht mehr erkannt wurden. Hier ist also ein gewisser Bias zu bedenken.

Der minimale C-Wert hat keine diagnostische Bedeutung und wurde nur zur Dokumentation aufgenommen. Gewisse Schwankungen zwischen beiden Messmethoden sind aus den im Ergebnisteil abgebildeten Diagrammen abzulesen; in den statistischen Kennwerten wie Mittelwert und Standardabweichung etc. ergaben sich aber keine wesentlichen Unterschiede.

Der maximale C-Wert hingegen ist gerade im Kontext der bei der Schilddrüse zu beachtenden euploiden Polyploidisierung (siehe oben) und der Berücksichtigung hoher Einzelwerte in der Diagnose einer Einzelzellaneuploidie von Bedeutung, ebenso wie unten der 9cEE. Auch hier weisen die statistischen Kennwerte kaum Unterschiede zwischen den Messreihen auf: Bei der manuellen Messung lag der maximale C-Wert zwischen 3,71c und 17,49c, bei der automatischen Messung zwischen 3,75c und 17,60c,

was auf eine sehr gute Vergleichbarkeit deutet. Auch die Mittelwerte waren mit 8,34c und 7,89c sehr ähnlich. Extrem hohe Maximalwerte, wie sie Heide an der Schilddrüse mit bis zu 128c beschrieben hatte (Heide, 1982), traten nicht auf. Andererseits zeigen das Streudiagramm und der Bland-Altman-Plot (Abb. 25 und 26), wie von Fall zu Fall eben doch bedeutendere Schwankungen vorkommen können.

Werte über einem bestimmten Level kann man in der DNA-Bildzytometrie als "EE" oder "ER" zahlentechnisch abbilden, wobei "ER" definitionsgemäß dem "EE in %" entspricht. Dass die Zahl der 5cEE und der 9cEE bei der automatischen Messung im Durchschnitt höher sein muss als bei der manuellen, ist einsichtig. Die Daten in Tabelle 14 und Tabelle 16 weisen diese Unterschiede auch aus; so ist beispielsweise der Mittelwert für den 5cEE bei der automatischen Messung fast genau 1,5-mal so hoch wie bei der manuellen (15,2 vs. 10,23). Gleiches gilt auch für den 9cEE mit einem Faktor von 1,98, wobei dort aber beide Durchschnittswerte gering waren (1,05 vs. 0,63). Andererseits zeigen die Abbildungen 26 und 27 sowie 32 und 33, dass nicht wenige Fälle vorkommen, bei denen die Werte für den 5cEE oder den 9cEE bei manueller Messung höher als bei der automatischen Messung waren - dieses betrifft alle Datenpunkte, die unterhalb der Winkelhalbierenden in den Streudiagrammen der Abbildungen 25 und 32 liegen. Dies bedeutet, dass es im Grunde keine Regelmäßigkeit gibt, die eine höhere Zahl von 5cEE oder 9cEE bei automatischen Messungen sicher voraussetzen lässt.

Eine Überlegung wäre daher, ob ein Ausweichen auf relative Maße, wie sie der 5cER und der 9cER bieten, nicht besser wäre. Die zugehörigen Daten zeigen hier im Mittel allerdings höhere Werte bei den manuellen Messungen, was wohl daran liegt, dass ein Anreichern der an sich randomisiert gedachten manuellen Messungen durch hohe Einzelwerte in der Routinediagnostik erlaubt war - denn das Auffinden solcher Zellen beeinflusst nicht die Lage der DNA-Stammlinie, verfälscht also nicht die Diagnostik bezogen auf die DNA-Stammlinienaneuploidie, macht aber ein Hinzunehmen des Kriteriums der Einzelzellaneuploidie möglich. Hier ist somit eine wirkliche Vergleichbarkeit der Daten aus messmethodischen Gründen nicht gegeben.

Die DNA-Stammlinie wiederum zeigte ebenfalls eine gewisse Variation zwischen den beiden Messmethoden. Die Werte bewegten sich bei der manuellen Messung zwischen 1,80c und 3,63c und bei der automatischen Messung zwischen 1,51c und 4,01c. Auffällig war, dass unbeschadet eines sehr hohen Korrelationskoeffizienten von $r = 0,896$ die meisten Zahlenpaare einen geringeren Messwert in der automatischen Messung aufwiesen, graphisch daran erkennbar, dass die meisten Punkte der Abbildung 36 unterhalb der Winkelhalbierenden lagen. Hier wäre möglicherweise die Einführung eines Korrekturfaktors zu diskutieren, um Fehlinterpretationen vermeiden zu helfen.

Zusammenfassend lässt sich sagen, dass durchaus eine gewisse Vergleichbarkeit in den Ergebnissen der beiden Messreihen besteht, dass sich aber auch nicht unwesentliche Unterschiede ergeben haben, die gerade in Einzelfällen die Ergebnisinterpretation

erschweren können. Im Folgenden wird darauf eingegangen, wie sich die Vergleichbarkeit in der diagnostischen Interpretation darstellt.

4.2.2 Vergleichbarkeit der diagnostischen Interpretation der manuellen und der automatischen DNA-Bildzytometrie

Insgesamt wurden bezogen auf die DNA-Stammlinienploidie 32 der 40 Fälle übereinstimmend als diploid oder aneuploid interpretiert und 8 divergent (20%), was in der Kappa-Statistik zu einer mittleren Übereinstimmung führte ($\kappa = 0,584$, "moderate"). Innerhalb der 8 divergenten Fälle waren die Abweichungen etwa gleich verteilt; in drei Fällen wies die manuelle Messung eine aneuploide Stammlinie auf und die automatische Messung eine euploide, in fünf Fällen war es umgekehrt.

Die diagnostische Interpretation des 9cEE war in 34 Fällen übereinstimmend und in 6 Fällen divergent (15%), weswegen sich in der Kappa-Statistik hier ein höherer Wert ergab ($\kappa = 0,60$, "moderate"). Jeweils drei Fälle mit einem oder mehreren 9cEE in der manuellen Messung zeigten in der automatischen Messung keine solchen Zellen und umgekehrt.

Für die Kombination beider diagnostischer Parameter ergab sich dann eine Übereinstimmungsrate von 29 Fällen bei 11 divergenten Beurteilungen (27,5%, $\kappa = 0,442$, "moderate"). Auch hier kamen etwa gleich viele Abweichungen in beide Richtungen vor. Sechs in der manuellen Messung euploide Fälle waren in der automatischen Messung aneuploid, in fünf Fällen fand sich ein umgekehrter Effekt.

Insgesamt zeigte sich somit, dass die größte Übereinstimmung in der Zuordnung von Aneuploidie und Euploidie mit 85% für den 9cEE möglich war. Die Lage der DNA-Stammlinien (Übereinstimmung 75%) und die Kombination aus 9cEE und Lage der DNA-Stammlinie (Übereinstimmung 72,5%) wiesen etwa ein identisches Resultat auf.

Insbesondere auffällig und problematisch ist in diesem Kontext die Tatsache, dass nicht alle Zellkerne mit DNA-Gehalten von $> 9c$ aus den manuellen Messreihen automatisch ebenfalls nachgewiesen werden konnten. Im Einzelfall wäre dieses dadurch erklärbar, dass möglicherweise einmal in der manuellen Messung eine Kernüberlappung nicht bemerkt worden wäre oder in der automatischen Messung ein etwas größerer Zellkern fälschlich als Artefakt interpretiert worden wäre. Bezogen auf die DNA-Stammlinie überraschte ebenfalls die relative hohe Zahl an Abweichungen, die an der Eignung des seinerzeit entwickelten *Classifiers* für die Schilddrüsenzytologie zweifeln lassen könnten.

Auf Details bezüglich der diskrepanten Fälle wird in Kapitel 4.3 eingegangen.

4.3 Interpretation der 14 zwischen manueller und automatischer Messung diskrepanten Fälle

Im Folgenden soll versucht werden, die Unterschiede in der diagnostischen Interpretation der 14 diskrepanten Fällen zu erklären. Dazu bietet sich an, die Fälle in drei unterschiedliche Gruppen zu gliedern.

4.3.1 Sechs Fälle mit Unterschieden in der Interpretation des 9cEE

Bei den sechs Fällen mit den Nummern 3, 4, 9, 31, 34 und 36, in denen bei nur einer beiden Messungen 9cEEs nachgewiesen worden waren, fällt auf, dass diese Messung jeweils nur genau eine Zelle über 9c aufwies. Die Fälle 3, 31 und 36 zeigten jeweils eine diploide und Fall 9 jeweils eine tetraploide DNA-Stammlinie und wären in diesem Auswertemodus als euploid interpretiert worden, die Fälle 4 und 34 waren in beiden Messungen aneuploid.

Zum einen wäre zu prüfen, ob die 9cEEs überhaupt als solche glaubhaft waren - insbesondere, wenn manuell nachgewiesene Zellen in der automatischen Messung, die ja theoretisch das gesamte Präparat erfassen soll, nicht nachweisbar waren. Leider lag in den Fällen 4, 31 und 36, bei denen diese Konstellation zutraf, keine Bildgalerie der manuellen Messung vor, die zum Zeitpunkt der Messung nicht obligatorisch abgespeichert wurde. Es lässt sich somit nicht überprüfen, ob hier möglicherweise ein Artefakt als Zellkern akzeptiert worden sein könnte oder ob Kernüberlagerungen für den Messwert verantwortlich sein könnten. In den übrigen drei Fällen 3, 9 und 34 wurden in der manuellen Messung keine 9cEEs nachgewiesen, in der automatischen Messung jeweils einer mit Messwerten von 10,58c, 9,11c bzw. 10,12c.

Zum anderen bleibt zu fragen, inwieweit man sich wirklich, gerade bei polyploidisierendem Gewebe, auf ein Kriterium wie den 9cEE stützen sollte, wenn man als Entscheidungsmerkmal dessen reine Existenz nimmt ($9cEE = 0$ versus $9cEE > 0$). Das kann gerade bei automatischen Messungen mit ihrer wesentlich höheren Zahl von Analysezellen zu Problemen führen - schließlich ist das Kriterium " $9cEE > 0$ " nur ein Erfahrungswert aus multiplen Messungen an ca. 300 Analysezellen. Misst man nun 3000 oder noch mehr Zellen, so ist das Risiko von falsch positiven Diagnosen zu bedenken. Eine mögliche Lösung wäre der Rückgriff auf den 9cER, den prozentualen Anteil der 9cEE, z.B. mit einem Kriterium für Einzelzellaneuploidie von $9cEE \geq 0,33\%$, was dann einem $9cEE \geq 1$, normiert auf 300 Analysezellen, entspräche. Diese Grenzziehung ist allerdings rein hypothetisch und nicht wissenschaftlich überprüft.

4.3.2 Fünf Fälle, in denen die in einer der beiden Messungen ermittelte aneuploide DNA-Stammlinie in der anderen Messung nicht bestätigt werden konnte

Manuelle Messung aneuploid, automatische Messung euploid

In den drei Fällen 10, 23 und 38 lagen bei der manuellen Messung aneuploide DNA-Stammlinien vor, während die automatischen Messungen euploide DNA-Stammlinien erbrachten. Alle drei Fälle waren für den 9cEE mit beiden Messverfahren negativ.

Bei Fall 10 mit der späteren histologischen Diagnose eines follikulären Adenoms vom oxyphilen Typ bei Struma colloidosa nodosa mit regressiven Veränderungen wurde manuell eine Stammlinie bei 2,69c dargestellt, während sich in der automatischen Messung eine Stammlinie bei 1,99c mit einem kleinen Verdopplungsgipfel bei 3,99c zeigte (Abb. 38). Da das Präparat teilweise sehr helle und somit durch den Screener schlecht erkennbare Bereiche aufwies, wurde dieses Präparat ein zweites Mal gemessen; dabei ergab sich ein gleichartiges Ergebnis. Vom DNA-Histogramm her hat man bei der automatischen Messung den Eindruck, als ob sich im Bereich oberhalb des Peaks bei 1,99c eine Art Schulter bildet, in der möglicherweise die Stammlinie bei 2,69c untergehen könnte; jedenfalls ist die Messwertverteilung rechtslastig und nicht nach Art einer Gauß-Verteilung symmetrisch um den Peak angeordnet. Die Abweichung in der Interpretation könnte somit damit erklärt werden, dass viele diploide Zellen, die nicht zur eigentlichen Läsion gehören, mitgemessen worden sein dürften.

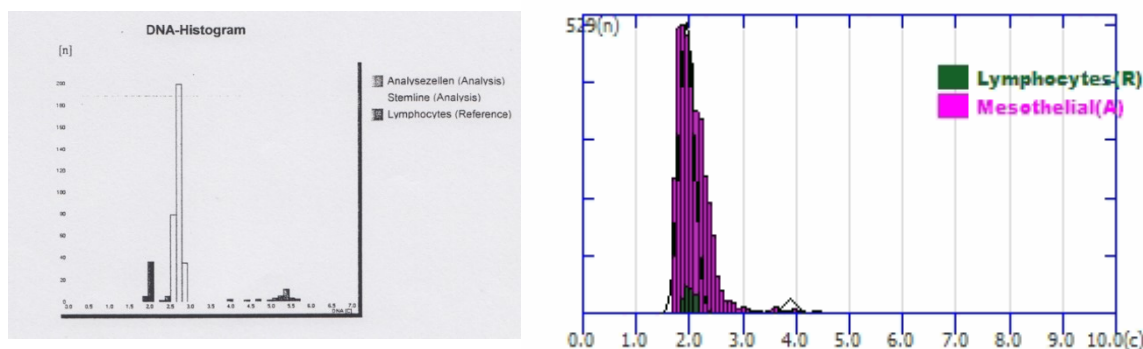


Abbildung 38: Histogramme der DNA-bildzytometrischen Messungen zu Fall 10. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.

Genauso verhält es sich mit Fall 38, bei welchem histologisch ein mikrofollikuläres Adenom mit regressiven Veränderungen diagnostiziert wurde. Auch hier zeigte sich bei der automatischen Messung ein asymmetrisches DNA-Histogramm, das ähnlich wie bei Fall 10 eine Schulterbildung in dem Bereich erkennen lässt, in dem bei der manuellen Messung die aneuploide DNA-Stammlinie identifiziert worden war. Auch hier kann man postulieren, dass durch die Untermischung mit normalen Schilddrüsenepithelien die aneuploide DNA-Stammlinie mengenmäßig untergegangen ist.

Anders sieht es bei Fall 23 aus, bei dem histologisch ein regressiv-zystisch verändertes follikuläres Adenom bei Kolloidstruma diagnostiziert wurde. Bei der manuellen

Messung hatte sich eine aneuploide DNA-Stammlinie bei 2,36c ergeben, bei der automatisch eine euploide Stammlinie bei 1,83c (Abb. 39). Aus dem DNA-Histogramm lässt sich der Unterschied in der Lage der DNA-Stammlinie, manuell bei 2,36c (aneuploid) und automatisch bei 1,83c (euploid), nicht auflösen. Beide DNA-Histogramme erscheinen schlank, im Peakbereich symmetrisch ohne Schulterbildungen und sind in sich jeweils überzeugend. Das Vorkommen von Referenzzellen genau bei 2c unterstreicht die Authentizität der jeweiligen Messung. Ein Grund für den Unterschied in den Messergebnissen lässt sich somit hier nicht herausarbeiten.

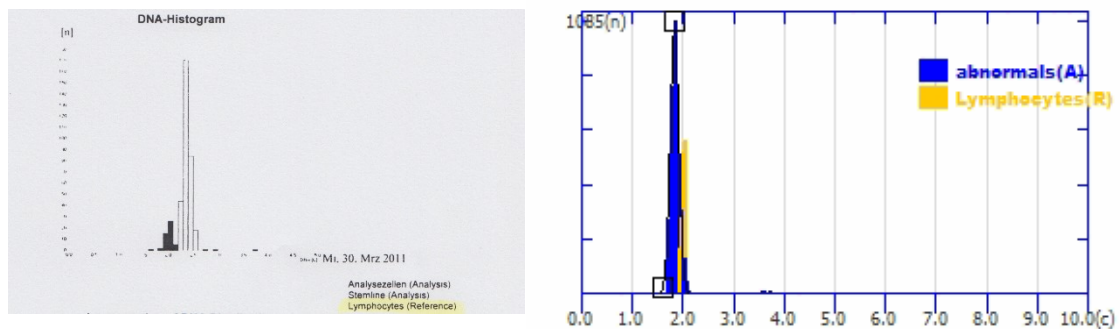


Abbildung 39: Histogramme der DNA-bildzytometrischen Messungen zu Fall 23. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.

Manuelle Messung euploid, automatische Messung aneuploid

Die Fälle 7 und 28 zeigten manuell eine euploide DNA-Stammlinie bei 2,06c bzw. bei 1,99c und bei der automatischen Messung aneuploide Werte bei 2,36c und bei 1,79c.

Bei Fall 7 lag trotz der zytologischen Diagnose eines onkozytären Adenoms letztendlich im Resektat lediglich eine nekrotische Knotenstruma vor. Der Vergleich der DNA-Histogramme zeigt in der manuellen Messung eine überzeugende DNA-Verteilung mit einer Stammlinie bei 2,06c und einem kleinen Verdopplungsgipfel bei 4c sowie auch bei 8c; die automatische Messung weist aber neben einem ebenfalls überzeugenden Peak bei 2c, den man wohl normalen Zellen zuordnen muss, eine aneuploide DNA-Stammlinie bei 2,36c aus, die schlank und relativ hoch und somit zweifelsfrei glaubhaft ist (Abb. 40). Hier liegt also in umgekehrter Richtung zum oben dargestellten Fall 23 eine nicht auflösbare Diskrepanz der Ergebnisse vor.

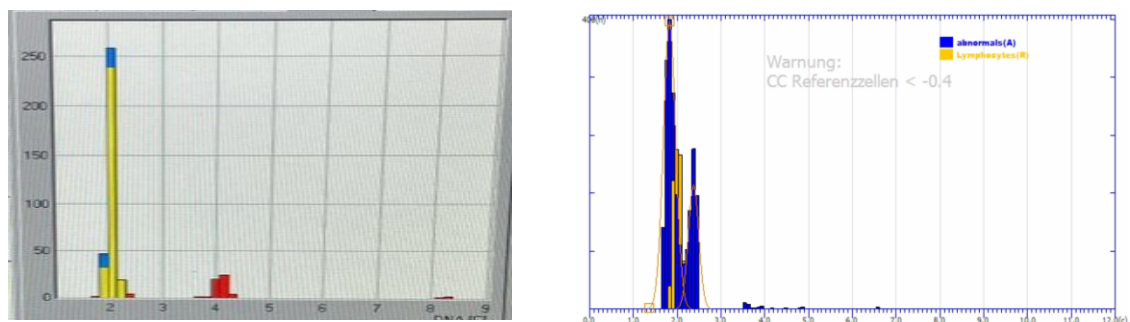


Abbildung 40: Histogramme der DNA-bildzytometrischen Messungen zu Fall 7. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.

Bei Fall 28 schließlich, der zytologisch als Thyreoiditis mit follikulärer Hyperplasie interpretiert worden war und histologisch gutartige entzündliche und fibrosierende Veränderungen gezeigt hatte, zeigten beide DNA-Histogramme ein identisches Grundmuster, jeweils mit einer hohen schlanken Peak-Bildung mit einer Verdopplung und Vervierfachung in entsprechenden Bereichen. Der Unterschied liegt darin, dass in der manuellen Messung die DNA-Stammlinie in den diploiden Bereich fiel (1,99c), während sie bei der automatischen Messung um ein Hunderstel unter der Grenze zum diploiden Bereich von 1,80c, nämlich bei 1,79c, lag - hier musste dann formal auf DNA-Aneuploidie entschieden werden. Sofern man eine solche Messung klinisch interpretieren möchte, wäre man aber gut beraten, sie in gewisser Weise relativiert in den Gesamtzusammenhang zu stellen. Hier wäre eine artefizielle Verschiebung der Daten in den hypodiploiden Bereich zu diskutieren.

4.3.3 Drei Fälle mit hypodiploiden automatischen Messergebnissen, die man für eine klinische Interpretation nicht akzeptieren dürfte

Schließlich bleiben noch drei Fälle, die in der Studie belassen wurden, weil die Präparate im automatischen Messvorgang noch technisch messbar gewesen waren, bei denen aber Ergebnissen ermittelt wurden, die nicht glaubhaft sind. Es handelt sich um die Fälle 14, 17 und 25.

Bei Fall 14 als Beispiel, der in der späteren histologischen Untersuchung eine regressiv veränderte Struma colloidosa nodosa mit zwei großen regressiv, follikulären Adenomen zeigte, ergab sich in der manuellen Messung ein klassisches, stark ausgeprägtes Bild einer euploiden Polyploidisierung mit Peaks und Verdopplungsgipfeln bei 2c, 4c und 8c und Zellen bis mindestens in den 16c-Bereich (Abb. 41). Die automatische Messung hingegen wies eine DNA-Verteilung auf, die einen als aneuploid zu benennenden Hauptpeak bei 1,51c mit kleinen angedeuteten Verdopplungen bei 3c und 6c zeigte. Außerdem lagen Einzelzellen oberhalb von 9c vor (9cEE = 6), die im Histogramm aufgrund ihrer geringer Zahl aber untergehen.

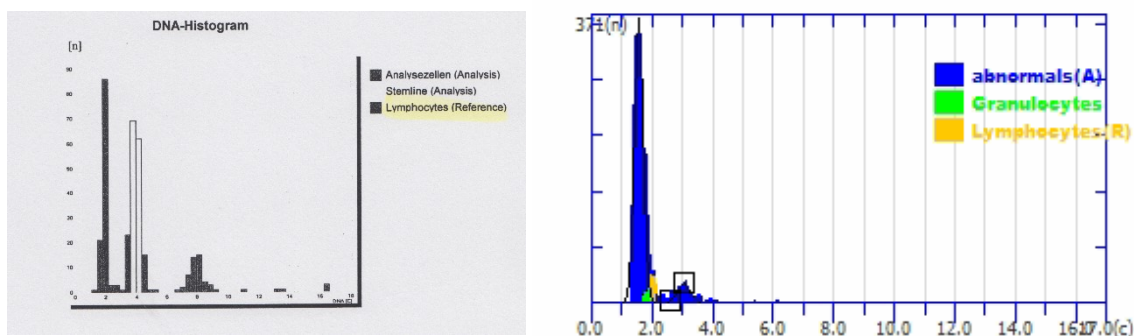


Abbildung 41: Histogramme der DNA-bildzytometrischen Messungen zu Fall 14. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.

Die Datenverteilung in der manuellen Messung ist mit Peaks und Verdopplungsgipfeln bei 2c, 4c und 8c und Zellen bis mindestens in den 16c-Bereich ohne weiteres glaubwürdig, so dass man die Interpretation als eine klassische DNA-Verteilung bei euploider Polyploidisierung auch retrospektiv aufrechterhalten kann. Die automatische Messung mit einem ersten Peak bei 1,51c und kleinen angedeuteten Verdopplungen bei 3c und 6c zeigt zwar ein gleichartiges mehrfaches Verdopplungsverhalten wie die manuelle Messung, ist aber mit ihrem Peak statt im diploiden im hypodiploiden Bereich angesiedelt. Hier gehen wir davon aus, dass bei der langjährigen Lagerung des Präparates möglicherweise Ausbleicheffekte, die die größeren Epithelzellen mehr betroffen haben als die kleineren Eichzellen, eine Rolle spielen.

Bei den Fällen 17 und 25 liegt jeweils ein ähnliches Problem wie bei Fall 14 vor. Hier war mit 1,83c bzw. 1,89c schon der Peak der manuellen Messung etwas nach links verschoben, aber noch im diploiden Bereich (1,8c - 2,2c), wohingegen in der automatischen Messung nur noch 1,58c bzw. 1,71c und somit hypodiploide Wert für die Lage der Stammlinie ermittelt wurden. Auch hier stellte sich die Frage nach der Glaubwürdigkeit einer hypodiploiden DNA-Stammlinie, von der wir hier ebenfalls nicht ausgehen können.

Insgesamt sollten, solange nicht zytogenetisch hypodiploide Chromosomensätze bei Tumorentitäten der Schilddrüse nachgewiesen worden sind, hypodiploide DNA-Stammlinien nicht akzeptiert werden, sondern als Messfehler klassifiziert werden, in unseren drei Fällen offenbar durch die längere Lagerung der Präparate mit zu postulierender Alterung und Ausbleichung bedingt. Ob man dann berechtigt wäre, die Lage der Stammlinie durch virtuelles Verschieben in den diploiden Bereich zu korrigieren, müsste an einem großen Kollektiv von ähnlich gelagerten Messwertvergleichen gezielter untersucht werden.

4.3.4 Zusammenfassende Beurteilung

Fasst man die 14 diskrepanten Fälle interpretatorisch zusammen, so ergeben sich

- sechs Fälle, bei denen die Interpretation vom Nachweis bzw. Nicht-Nachweis einer singulären Zelle mit einem DNA-Gehalt von $> 9c$ abhing (Fälle 3, 4, 9, 31, 34 und 36),
- vier Fälle, in denen die Diskrepanz wohl auf eine unterschiedliche Alterung von Referenz- und Analysezellen zurückgeht und zu einem artefiziellen Abrutschen der Analysezellen in den hypodiploiden Bereich geführt hat (Fälle 14, 17, 25 und 28),
- zwei Fälle, in denen man davon ausgehen kann, dass in der automatischen Messung eine kleine aneuploide DNA-Stammlinie durch eine viele größere

diploide DNA-Stammlinie so stark überdeckt wurde, dass sie sozusagen "durchgerutscht" ist und auf diese Weise automatisch ein falsch-negativer euploider Befund erhoben wurde (Fälle 10 und 38),

- zwei Fälle, in denen sich die Diskrepanzen zwischen der Lage der DNA-Stammlinie in der manuellen und der automatischen Messung nicht aufklären lassen (Fälle 7 und 23).

Aus den Ergebnissen sollte man ableiten,

- dass sich für zwei der diskrepanten Fälle (7 und 23) keine Erklärungen finden lassen und somit der DNA-bildzytometrische Charakter der untersuchten Zellpopulation offen bleiben muss,
- dass sich die Unterschiede in vier Fällen (14, 18, 25 und 28) in erster Linie aus Gründen der Präparatealterung erklären, also in ihrer Diskrepanz methodischer Natur sind und
- dass sich zu den übrigen acht diskrepanten Fällen inhaltliche Überlegungen anschließen müssen, wie man die Interpretation von Messungen modifizieren und verbessern kann. Insbesondere wäre zu fragen, ob die Grenze für die DNA-Einzelzellaneuploidie von 0 für den 9cEE aufrechterhalten werden kann oder ob ein Bezug auf den 9cER adäquater wäre, gerade wenn die Zahl von Analysezellen hoch ist (3, 4, 9, 31, 34 und 36). Auch die beiden Fälle, in denen kleine Stammlinien bei hoher Zahl von Analysezellen untergehen, fallen in diesen Bereich hinein (10, 38) - ähnliche Probleme sind, ohne jetzt hierauf näher eingehen zu wollen, auch in der DNA-Durchflusszytometrie, in der es um noch wesentlich höhere Zahlen von Analysezellen geht, bekannt.

4.4 Diagnostische Treffsicherheit der manuellen und der automatischen DNA-Bildzytometrie in der Schilddrüsenzytologie

Was soll die DNA-Bildzytometrie in der präoperativen zytologischen Diagnostik leisten? Wenn es in der Schilddrüsenzytologie darum geht, bei der Entscheidung zu einer Operation einen Beitrag zu leisten, so wäre das Ziel, die Fälle zu identifizieren, die wirklich negativ sind, also das geringste tumorbiologische Risiko tragen. Entsprechend wäre bei der Diskussion der Ergebnisse in erster Linie der Vergleich "negativ" gegen "zweifelhaft / dringender Verdacht / positiv" zu betrachten. Dabei muss allerdings beachtet werden, dass sich damit in beiden Gruppen ein Teil der "follikulären Läsionen" (Struma, follikuläres Adenom, follikuläres Karzinom) wiederfinden werden. Denn wie in Kapitel 1.3.5 ausgeführt, sind sowohl negative Zytologien, als auch zweifelhafte Zytologien Korrelate für Strumen unterschiedlicher Zellularität.

Tabelle 37 fasst noch einmal die Ergebnisse der Tabellen 22, 27 und 32 zusammen.

Zytologische Diagnose: "negativ" gegen "zw. / dV / pos."	Auswertemodus					
	DNA-Stammlinie		9cEE		Kombinierte Auswertung	
	manuell	autom.	manuell	autom.	manuell	autom.
Sensitivität	39,4%	48,5%	30,3%	24,2%	60,6%	60,6%
Spezifität	71,4%	85,7%	100%	71,4%	71,4%	57,1%
PPV	86,7%	94,1%	100%	80,0%	90,9%	87,0%
NPV	20,0%	26,1%	23,3%	16,7%	27,8%	23,5%
Gesamttrefferlichkeit	45,0%	50,0%	42,5%	32,5%	62,5%	60,0%

Tabelle 37: Treffsicherheit der DNA-Bildzytometrie nach Lage der DNA-Stammlinie, des 9cEE und der kombinierten Auswertung, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv

Es zeigt sich, dass man mit der DNA-Bildzytometrie, bezogen auf die zytologische Diagnostik als Referenz, zwar eine vertretbare Spezifität erreichen kann, dass aber die Sensitivität recht gering ist. Dies bedeutet, dass man bei der Schilddrüsenzytologie keinen allgemeinen Zugewinn an Informationen durch die DNA-Bildzytometrie erhält - was aber nicht für den Einzelfall und die dort gestellte Indikation zur Untersuchung gelten muss.

Außerdem ergab sich kein eindeutiger Trend, ob die automatische DNA-Bildzytometrie zu besseren Ergebnissen als die manuelle führt. Die Automatisierung der Methode zum inzwischen sieben Jahren alten Software-Stand bei der Erstellung der Messungen erbringt somit bezogen auf die zytologische Diagnostik keinen allgemeinen Vorteil. Allerdings kann man, wie unten auch bei der Besprechung der diskrepanten Fälle ersichtlich, die Messreihe auch als eine Studie des Lernens ansehen, aus der sich Schlüsse für weitere Planungen der Automatisierung bzw. der Interpretation ihrer Ergebnisse ziehen lassen.

Im Hinblick auf Ergebnisse einer späteren Schilddrüsenhistologie nach Resektion wäre wünschenswert, wenn sich mit der Zytologie oder einer adjuvanten Methode die Dignität eines Tumors vorhersagen ließe oder wenigstens geklärt werden könnte, ob ein Tumor oder eine reaktive Veränderung vorliegt. Dieses ist zytologisch zwar der Fall bei papillären, medullären und anaplastischen Karzinomen, gilt aber erneut nicht für die sehr viel häufigeren follikulären Läsionen.

Tabelle 38 gibt hierzu die Ergebnisse der Tabellen 24, 29 und 34 zusammengefasst wieder.

Histologische Diagnose: "kein Tumor" gegen "Tumor"	Auswertemodus					
	DNA-Stammlinie		9cEE		Kombinierte Auswertung	
	manuell	autom.	manuell	autom.	manuell	autom.
Sensitivität	60,0%	60,0%	28,0%	28,0%	76,0%	72,0%
Spezifität	100%	86,7%	80,0%	80,0%	80,0%	66,7%
PPV	100%	88,2%	70,0%	70,0%	86,4%	78,2%
NPV	60,0%	56,5%	40,0%	40,0%	66,7%	58,8%
Gesamttrefferlichkeit	75,0%	70,0%	47,5%	47,5%	77,5%	70,0%

Tabelle 38: Treffsicherheit der DNA-Bildzytometrie nach Lage der DNA-Stammlinie, des 9cEE und der kombinierten Auswertung, bezogen auf das histologische Ergebnis

Es fällt auf, dass deutliche Unterschiede zwischen den Ergebnissen für die DNA-Stammlinie und den 9cEE bestehen. Insgesamt ergaben sich mit der kombinierten Auswertung die günstigsten Zahlenkonstellationen, wobei die manuelle Auswertung noch etwas bessere Zahlenwerte erbrachte.

4.5 Vergleich der Ergebnisse mit der Literatur

Die Treffsicherheit der manuellen DNA-Bildzytometrie wurde in der Literatur bereits für zahlreiche Organe untersucht und darf insgesamt als gesichert gelten. Auch die klinischen Erfahrungen am Standort Düsseldorf, wo die Methode in der klinischen Diagnostik regelmäßig eingesetzt wird, zeugen davon. Insofern kann, ohne dass hier die genauen Daten erörtert werden können, die manuelle DNA-Bildzytometrie als Referenz und Standard gesetzt werden.

Ihr Einsatzgebiet liegt allerdings kaum in der Abklärung unklarer Läsionen der Schilddrüse, sondern vielmehr in Anwendungen an Ergüssen oder an Gallenwegsbürstungen oder Pankreaspunktaten mit dem Ziel, über den Nachweis von DNA-Aneuploidie eine maligne Transformation der untersuchten Zellpopulation nachzuweisen. Entsprechend wären auch Präparationen von Entnahmen aus diesen Bereichen besondere Kandidaten für die Automatisierung der Methode. Bei der Entwicklung der automatischen DNA-Bildzytometrie wurde darauf auch Bezug genommen und gerade der für Automatisierung notwendige *Classifier* für Ergüsse als einer der ersten entwickelt.

In dieser Arbeit wurde dieser *Classifier* im Grunde organfremd an Schilddrüsenzytologien angewandt. Diese Vorgehensweise erschien vertretbar, da sich Zellkerne von Mesothelien und von Schilddrüsenepithelien in ihren morphologischen Eigenschaften von vornherein ähneln und im Besonderen auch bei Bildung tumorartiger oder wirklich tumoröser Veränderungen oft keine klassischen Atypien aufweisen -

dieses führt dazu, dass manche Probleme der Schilddrüsenzytologie sich in gewisser Ähnlichkeit auch in der Ergusszytologie zeigen.

Es macht daher Sinn, die Ergebnisse dieser Studie in erster Linie mit denjenigen zu vergleichen, die aus unserer Arbeitsgruppe von Feider (Dissertation 2018) bzw. Böcking et al. (Publikation hierzu 2019) an Ergusspräparaten vorgelegt wurden. Dabei soll wegen der unterschiedlichen spezifischen zytologischen Fragestellungen nur auf methodische Aspekte eingegangen werden.

Bei Feider (2018) lassen sich folgende wichtigen Kenndaten aus den an 121 Fällen stammenden Messungen extrahieren:

- 81 Fälle zeigten mit beiden Messverfahren eine diploide DNA-Stammlinie. In sieben Fällen lag eine DNA-Stammlinienaneuploidie nur in der manuellen Messung vor, in sechs Fällen nur automatisch, in den weiteren 27 Fällen mit beiden Verfahren. Der Anteil diskrepanter Fälle errechnet sich somit mit 10,7% (13/121).
- 87 Fälle wiesen mit beiden Messverfahren keine 9cEE auf; in zwei Fällen wurden sie nur in der manuellen Messung nachgewiesen, in sechs Fällen nur automatisch, in den übrigen 26 mit beiden Verfahren. Diskrepanz beurteilt wurde somit die Zahl der 9cEE in 6,6% der Fälle (8/121). In der automatischen Messreihe war die Zahl der 9cEE im Mittelwert 1,8-mal so hoch wie manuell.
- 79 Fälle wurden in der kombinierten Auswertung als jeweils euploid eingestuft, 37 als jeweils aneuploid. Fünf Fälle waren in der automatischen Messung aneuploid, in der manuellen euploid; die gegenteilige Konstellation kam nicht vor. Abschließend diskrepanz diagnostiziert wurden somit 4,1% der Fälle (5/121).

In der eigenen Studie lagen die Abweichungen zwischen den Messverfahren deutlich höher, nämlich bei 20% für die DNA-Stammlinie (8/40), bei 15% für den 9cEE (6/40) und bei 27,5% für die kombinierte Auswertung (11/40).

Zu fragen wäre nach Gründen für diese höheren Abweichungen.

Zunächst wären Unterschiede in der Präparation zu diskutieren. Die Präparate unserer Studie waren seitens der punktierenden Ärzte bereits als fertige Ausstriche, nicht hingegen als Flüssigkeiten, eingesandt worden. Diese wiesen zum Teil massive Überlagerungen zwischen mehreren kernhaltigen Zellen auf, aber auch zwischen einer kernhaltigen Zelle und Erythrozyten, was die Erkennung der Messzellen durch das Bildverarbeitungssystem erschwerte. In der Studie von Feider hingegen wurden in der Cytopathologie selber hergestellte Sedimentausstriche zur Messung verwendet, die von

erfahrenen MTAs angefertigt worden waren und somit sicherlich annähernd durchgehend Monolayer-Qualitäten aufwiesen. Insofern ist in unserer Studie das Risiko von (morphologisch nicht immer erkennbaren) Kernüberlagerungen höher gewesen. Diese können zwar nicht zu einer falschen DNA-Stammlinie führen, wären aber als Ursache für falsche 9cEE denkbar. Diverse solcher Ereignisse wurden auch als solche ebenfalls bei der Durchsicht der Image-Galerien erkannt und eliminiert.

Außerdem waren die Fälle bei uns zum Teil bereits viele Jahre zuvor Feulgen-gefärbt worden, was das Risiko des Alterns der Präparate mit sich brachte, während die Präparate der Erguss-Studie aktuell als zusätzliche Ausstriche gewonnen worden waren. Letzteres war möglich, weil Ergüsse eben zur "Massenware" der zytologischen Diagnostik gehören und entsprechend häufig auch DNA-bildzytometrisch behandelt wurden, während Schilddrüsenzytologien inzwischen relativ selten im Krankengut vorkommen und noch seltener adjuvant untersucht wurden. Die in Kapitel 3.3.2 dargestellten Werte für die Referenzzell-IOD belegen zwar, dass die Referenzzell-IOD der automatischen Messungen im Mittel bei 94,75% des Wertes der manuellen Messungen lag (307,27 vs. 324,27). Es ist aber als Hypothese aufzustellen, dass die Ausbleicheffekte pro Zellart nicht gleichmäßig waren. Bei manchen automatischen Messungen traten hypodiploide DNA-Stammlinien auf, die eigentlich nicht glaubhaft sind, da dieses Phänomen an der Schilddrüse nicht auftritt, sondern eigentlich, auch zytogenetisch belegt, bei soliden Tumoren auf chromophobe Nierenzellkarzinome beschränkt ist (G. Li et al., 2005) (Alaghebandan et al., 2021). Diese Fälle wurden zwar belassen (siehe auch Kapitel 4.3), es bleibt aber zweifelhaft, ob sie wirklich reale Messungen repräsentieren. Würde man sie eliminieren, säne die Zahl der beurteilbaren Fälle von 40 auf 37, es ergäben sich dann aber geringere Abweichungen in den Kennwerten für die Diskrepanzen von DNA-Stammlinie, 9cEE und kombinierter Auswertung von dann noch 11 von 37 Fällen (29,7%) statt 14 von 40 (35,0%).

Selbstverständlich stellt sich auch die Frage nach der Eignung des *Classifiers*. Hier ergaben sich aber keine Hinweise hierfür, dass die Erkennung von Thyreozyten als Epithelzellen (in Abgrenzung gegenüber Artefakten, Granulozyten oder Entzündungszellen) nicht möglich sein sollte. Der *Classifier* wurde bei der Planung der Studie als passend empfunden, weil er für "ähnliche Zellen", nämlich Mesothelien, und für die gleiche Präparationsart, nämlich für luftgetrocknete Ausstriche bzw. Sedimentausstriche, konzipiert war.

Insgesamt zeigen die Vergleiche mit der Arbeit von Feider aber, dass bei der Automatisierung der DNA-Bildzytometrie bei Schilddrüsenausstrichen noch einiges zu tun bleibt, jedenfalls bezogen auf die Softwareversion, die ab 2014 im Einsatz war. Möglicherweise wäre mit einem eigenen *Classifier* für Schilddrüsenepithelien noch eine gewisse Verbesserung möglich; aus eigener Erfahrung mit den Messungen ist aber insbesondere die Ausstrichqualität, die oft kein Monolayer-Niveau erreichte, ein Problem. Hier könnte möglicherweise die Umstellung der klassischen Schilddrüsen-

zytologie auf Dünnschichtverfahren helfen, so wie sie am Universitätsklinikum Essen, das sehr viele Schilddrüsenzytologien beurteilt, längst vorgenommen wurde. Dieses aber würde andererseits die Erstellung eines eigenen Schilddrüsen-*Classifiers* zwingend notwendig machen, da die Morphologie von Thyreozytenkernen im Alkohol-fixierten Dünnschichtpräparat bezüglich Größe, Form, Chromatineigenschaften etc. nicht ohne weiteres mit der im Luft-getrockneten klassischen Ausstrichpräparat vergleichbar sein wird.

Einem ganz anderen Interpretationsansatz ging die Studie zur Dissertation von Rezaie (2018) in unserer Arbeitsgruppe nach, bei der es darum ging, gezielt Fälle von Prostatakarzinomen nachzumessen, deren DNA-bildzytometrisches Profil zuvor als Typ B bzw. als Typ C nach Tribukait beurteilt worden war (Tribukait, 1993). Bei dieser Auswertung geht es darum, Urologen und Onkologen über die Möglichkeiten und Grenzen einer *Active Surveillance*-Strategie beim Prostatakarzinom zu beraten; bei einem Typ B erscheint diese noch möglich, bei Typ C wird sie als nicht vertretbar angesehen. Ursprünglich waren 30 manuell bereits in der Krankenversorgung ausgewertete Fälle zur Untersuchung vorgegeben, von denen 27 sich noch für die automatische DNA-Bildzytometrie als messbar erwiesen. Zwölf von ihnen entsprachen manuell einem Typ Tribukait B, fünfzehn einem Typ Tribukait C. Die Schwankungsbreiten zwischen der Lage der DNA-Stammlinien war relativ gering; so lag die Differenz zwischen den diploiden Stammlinien bei Tribukait B-Fällen von den Werten der manuellen und der automatischen Messung zwischen $-0,02c$ und $0,09c$, bei Tribukait C zwischen $-0,13c$ und $0,15c$. In der automatischen Messreihe, in der alle Präparate zweimal gemessen wurden, erwiesen sich aber dennoch fünf Fälle als diskrepant (18,5%, 5/27). Unterschiedliche Ergebnisse innerhalb der automatischen Doppelmessungen kamen nicht vor.

Methodische Studien anderer Arbeitsgruppen zur automatischen DNA-Bildzytometrie liegen nicht vor. Mit klinischem Blickwinkel zu nennen wäre eine Studie aus China, bei der ihre klinische Treffsicherheit an 117 Ergusspräparaten untersucht wurde (Meng et al., 2013). Dabei wurde ein Schwellenwert von mehr als vier Zellen mit einem DNA-Gehalt von $> 2,5c$ (entsprechend $2,5cEE > 4$) als optimiertes Kriterium für Malignität ermittelt, bezogen auf 1000 Messzellen. In Anlehnung an dieses Kriterium erreichte die Sensitivität bei der automatischen DNA-Bildzytometrie 83,3%, die Spezifität 100%; ein methodischer Vergleich mit einer etwaigen früheren manuellen Messreihe wurde nicht angeboten. Außerdem wurden aktuell zwei Arbeiten mit Messsystemen, die unserem verwandt sind, vorgelegt, die mit der automatischen DNA-Bildzytometrie klinisch vielversprechende Ergebnisse zur Nutzbarkeit der Methode in der gynäkologischen Krebsvorsorge (Cang et al., 2020) und in der Beurteilung von Mundhöhlenabstrichen erzielt hatten (Parfenova et al., 2021). Methodische Vergleiche mit manuellen Messungen wurden auch dort nicht vorgenommen.

5 Literaturverzeichnis

- Abram, M., Huhtamella, R., Kalfert, D., Hakso-Mäkinen, H., Ludvíková, M., & Kholová, I. (2021). The role of cell blocks and immunohistochemistry in thyroid atypia of undetermined significance / follicular lesion of undetermined significance Bethesda category. *Acta Cytologica*, 65(3), 257-263.
- Alaghehbandan, R., Trpkov, K., Tretiakova, M., Luis, A. S., Rogala, J. D., & Hes, O. (2021). Comprehensive review of numerical chromosomal aberrations in chromophobe renal cell carcinoma including its variant morphologies. *Advances in Anatomic Pathology*, 28(1), 8–20.
- Al-Brahim, N., & Asa, S. L. (2006). Papillary thyroid carcinoma: An overview. *Archives of Pathology & Laboratory Medicine*, 130(7), 1057–1062.
- Anonymus (2013). Endokrinologen: Viele Schilddrüsenoperationen sind überflüssig. Online unter <https://www.aerzteblatt.de/nachrichten/53696/Endokrinologen-Viele-Schilddruesenoperationen-sind-ueberfluessig>, abgerufen am 6.9.2021
- Asari R., Koperek O., Scheuba C., Riss P., Kaserer K., Hoffmann M., Niederle B. (2009). Follicular thyroid carcinoma in an iodine-replete endemic goiter region: A prospectively collected, retrospectively analyzed clinical trial. *Annals of Surgery* 249 (6), 1023-1031.
- Avetisian, I.L., & Petrova, G.V. (1996). Latent thyroid pathology in residents of Kiev, Ukraine. *Journal of Environmental Pathology, Toxicology and Oncology*, 15(2–4), 239–243.
- Bakhos, R., Selvaggi, S. M., DeJong, S., Gordon, D. L., Pitale, S. U., Herrmann, M., & Wojcik, E. M. (2000). Fine-needle aspiration of the thyroid: Rate and causes of cytohistopathologic discordance. *Diagnostic Cytopathology*, 23(4), 233–237.
- Bartolazzi, A., Orlandi, F., Saggiorato, E., Volante, M., Arecco, F., Rossetto, R., Palestini, N., Ghigo, E., Papotti, M., Bussolati, G., Martegani, M. P., Pantellini, F., Carpi, A., Giovagnoli, M. R., Monti, S., Toscano, V., Sciacchitano, S., Pennelli, G. M., Mian, C., Italian Thyroid Cancer Study Group (ITCSG). (2008). Galectin-3-expression analysis in the surgical selection of follicular thyroid nodules with indeterminate fine-needle aspiration cytology: A prospective multicentre study. *The Lancet Oncology*, 9(6), 543–549.
- Bartsch, D. K., Luster, M., Buhr, H. J., Lorenz, D., Germer, C.-T., Goretzki, P. E., & German Society for General and Visceral Surgery. (2018). Indications for the surgical management of benign goiter in adults. *Deutsches Ärzteblatt International*, 115(1–02), 1–7.
- Bergenfels, A., Jansson, S., Kristoffersson, A., Mårtensson, H., Reihner, E., Wallin, G., & Lausen, I. (2008). Complications to thyroid surgery: Results as reported in a database from a multicenter audit comprising 3,660 patients. *Langenbeck's Archives of Surgery*, 393(5), 667–673.
- Besic, N., Hocevar, M., Zgajnar, J., Pogacnik, A., Grazio-Frkovic, S., & Auersperg, M. (2005). Prognostic factors in anaplastic carcinoma of the thyroid-a multivariate

- survival analysis of 188 patients. *Langenbeck's Archives of Surgery*, 390(3), 203–208.
- Biesterfeld, S., Gerres, K., Fischer-Wein, G., & Böcking, A. (1994). Polyploidy in non-neoplastic tissues. *Journal of Clinical Pathology*, 47(1), 38–42.
- Bisi, H., Fernandes, V. S., de Camargo, R. Y., Koch, L., Abdo, A. H., & de Brito, T. (1989). The prevalence of unsuspected thyroid pathology in 300 sequential autopsies, with special reference to the incidental carcinoma. *Cancer*, 64(9), 1888–1893.
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1(8476), 307–310.
- Blank, W., & Braun, B. (2007). Sonografie der Schilddrüse. Teil 1 - Untersuchungstechnik, Normalbefund, Struma diffusa und Struma nodosa. *Ultraschall in der Medizin - European Journal of Ultrasound*, 28(6), 554–575.
- Böcking, A. (1998a). Abklärung plattenepithelialer Dysplasien mittels DNA-Bildzytometrie. *Deutsches Ärzteblatt*, 95(12), A-658 / B-570 / C-524.
- Böcking, A. (1998b). Standardisierte Befunderstellung in der extragenitalen Zytologie. *Der Pathologe*, 19 (3), 235-258.
- Böcking, A. (2002). Identifizierung progredienter Dysplasien des Plattenepithels mittels DNA-Bildzytometrie. *MTA Dialog* 6 (3), 490-493.
- Böcking, A., Giroud, F., & Reith, A. (1995). Consensus report of the ESACP task force on standardization of diagnostic DNA-image cytometry. *Anal Cell Pathol.*, 8(1), 67–74.
- Böcking, A., Friedrich, D., Meyer-Ebrecht, D., Zhu, C., Feider, A., & Biesterfeld, S. (2019). Automated detection of cancer cells in effusion specimens by DNA karyometry. *Cancer Cytopathology*, 127(1), 18–25.
- Böcking, A. (1995). DNA Measurements. When and Why? Compendium on quality assurance, proficiency testing and workload limitations in clinical cytology, *Tutorials of Cytology*, 170–188.
- Böcking, A., Chen, J., Friedrich, D., & Meyer-Ebrecht, D. (2015). Computerunterstützte Erkennung von Krebszellen mittels DNA-Karyometrie. *Trillium Krebsmedizin*, Heft 5.
- Böcking, A. & Nguyen, V.Q.H. (2004). Diagnostic and prognostic use of DNA image cytometry in cervical squamous intraepithelial lesions and invasive carcinoma. *Cancer*, 102(1), 41–54.
- Böcking, A., Striepecke, E., Auer, H., & Füzesi, L. (1994). Static DNA-cytometry. Biological background, technique and diagnostic interpretation. In: *Compendium on the Computerized Cytology and Histology Laboratory*. Wied G. L., Bartels P. H., Rosenthal D. L. (Eds.). *Tutorials of Cytology*, Chicago, Seite 107- 128.
- Böcking A., Tils, M., Schramm, M., Dietz, J., & Biesterfeld, S. (2014). DNA-cytometric grading of prostate cancer - systematic review of the literature with descriptive data analysis. *Pathology Discovery*, 2, 7

- Bramley, M. D., & Harrison, B. J. (1996). Papillary microcarcinoma of the thyroid gland. *The british journal of surgery*, 83(12), 1674–1683.
- Bruhn, H. D., Junker, R., Schäfer, H., & Schreiber, S. (2011). *LaborMedizin: Indikationen, Methodik und Laborwerte, Pathophysiologie und Klinik* (3. Auflage). Schattauer.
- Bubendorf, L., Feichter, G. E., Obermann, E. C., Dalquen, P. (2011a). Zytopathologie: Kapitel 3: Zytologische Tumorkriterien. In: Klöppel, G., Kreipe, H., Remmele, W. (Hrsg.): *Pathologie*. 3. Auflage, Springer Verlag, Berlin, S. 33-46.
- Bubendorf, L., Feichter, G. E., Obermann, E. C., Dalquen, P. (2011b). Zytopathologie. Kapitel 20: Schilddrüse. In: Klöppel, G., Kreipe, H. H., Remmele, W. (Hrsg.): *Pathologie*. 3. Auflage, Springer Verlag, Berlin, S. 429-458.
- Cai, Y.-F., Wang, Q.-X., Ni, C.-J., Guo, G.-L., Li, Q., Wang, O.-C., Wu, L., Du, H.-Y., You, J., & Zhang, X.-H. (2015). The clinical relevance of psammoma body and Hashimoto thyroiditis in papillary thyroid carcinoma. *Medicine*, 94(44), e1881.
- Cang, W., Li, Q., Gu, L., Hong, Z., Hu, Y., Di, W., & Qiu, L. (2020). Clinical evaluation of DNA ploidy for the triage of HPV-positive chinese women during cervical cancer screening. *Cancer Prevention Research*, 14 (3), 355-362.
- Carpi, A., Ferrari, E., Toni, M. G., Sagripanti, A., Nicolini, A., & Di Coscio, G. (1996). Needle aspiration techniques in preoperative selection of patients with thyroid nodules: A long-term study. *Journal of Clinical Oncology. Official Journal of the American Society of Clinical Oncology*, 14(5), 1704–1712.
- Castro, P., Sansonetty, F., Soares, P., Dias, A., & Sobrinho-Simões, M. (2001). Fetal adenomas and minimally invasive follicular carcinomas of the thyroid frequently display a triploid or near triploid DNA pattern. *Virchows Archiv: An International Journal of Pathology*, 438(4), 336–342.
- Chatelain, R., Willms A., Biesterfeld, S., Auffermann, W. & Böcking, A. (1989). Automated Feulgen staining with a temperature-controlled staining machine. *Analytical and Quantitative Cytology and Histology*, 11 (3), 211- 217.
- Cibas, E.S., & Ali, S.Z. (2017). The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid*, 27(11), 1341–1346.
- Cohen, J. (1960): A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20 (1), 37-46.
- Cooper, D. S., Doherty, G. M., Haugen, B. R., Hauger, B. R., Kloos, R. T., Lee, S. L., Mandel, S. J., Mazzaferri, E. L., McIver, B., Pacini, F., Schlumberger, M., Sherman, S. I., Steward, D. L., & Tuttle, R. M. (2009). Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*, 19(11), 1167–1214.
- Cooper, D.S., Doherty, G.M., Haugen, B.R., Kloos, R. T., Lee, S.L., Mandel, S.J., Mazzaferri, E.L., McIver, B., Sherman, S.I., Tuttle, R.M., & American thyroid association guidelines taskforce. (2006). Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*, 16(2), 109–142.

- Cotterill, S., Pearce, M., & Parker, L. (2001). Thyroid cancer in children and young adults in the north of England. Is increasing incidence related to the Chernobyl accident? *European Journal of Cancer*, 37(8), 1020–1026.
- DeLellis R. A., Lloyd R. V., Heitz P. U. & Eng, C. (Hrsg.) (2004). *Pathology and Genetics of Tumours of Endocrine Organs. WHO Classification of Tumours*, 3rd Edition, Volume 8, IARC Press, Lyon
- Dralle, H., & Machens, A. (2008). Surgical approaches in thyroid cancer and lymph-node metastases. *Best Practice & Research. Clinical Endocrinology & Metabolism*, 22(6), 971–987.
- Dralle, H., Sekulla, C., Haerting, J., Timmermann, W., Neumann, H. J., Kruse, E., Grond, S., Mühlig, H. P., Richter, C., Voss, J., Thomusch, O., Lippert, H., Gastinger, I., Brauckhoff, M., & Gimm, O. (2004). Risk factors of paralysis and functional outcome after recurrent laryngeal nerve monitoring in thyroid surgery. *Surgery*, 136(6), 1310–1322.
- Droese, M. (1979). *Aspirationszytologie der Schilddrüse*. F.K. Schattauer Verlag Stuttgart/New York.
- Droese, M. (1995). *Punktionszytologie der Schilddrüse, Atlas und Handbuch*. 2. Auflage. F.K. Schattauer Verlag Stuttgart/New York.
- Edmund, S., Cibas, M., Syed, Z., & Ali, M. (2009). The Bethesda system for reporting thyroid cytopathology. *American Journal of Clinical Pathology*, 132 (5), 658–665
- Farahati, J., Wegscheider, K., Christ, K., Gilman, E., & Oing, W. (2006). Gender-specific determinants of goiter. *Biological Trace Element Research*, 113(3), 223–230.
- Feider, A. (2018). *Erprobung eines automatischen DNA-zytometrischen Messsystems an luftgetrockneten Ergusspräparaten*. Med. dent. Diss., Heinrich-Heine-Universität Düsseldorf.
- Feldkamp, J., Führer, D., Luster, M., Musholt, T. J., Spitzweg C., & Schott, M. (2016). Fine needle aspiration in the investigation of thyroid nodules - indications, procedures and interpretation. *Deutsches Ärzteblatt International* 2016; 113: 353–359.
- Feulgen, R., & Rossenbeck, H. (1924). Mikroskopisch-chemischer Nachweis einer Nucleinsäure vom Typus der Thymonucleinsäure und die darauf beruhende elektive Färbung von Zellkernen in mikroskopischen Präparaten. *Z. Physiol. Chem.*, 135, 203–248.
- Francia, G., Azzolina, L., Mantovani, T., Daví, M. V., Brazzarola, P., Valerio, A., Petronio, R., & Sussi, P. (1997). Heterogeneity of nuclear DNA pattern and its relationship with cell cycle activity parameters in multinodular goitre. *Clinical Endocrinology*, 46(6), 649–654.
- Frates, M.C., Benson, C.B., Charboneau, J.W., Cibas, E.S., Clark, O.H., Coleman, B.G., Cronan, J.J., Doubilet, P.M., Evans, D.B., Goellner, J.R., Hay, I.D., Hertzberg, B.S., Intenzo, C.M., Jeffrey, R.B., Langer, J.E., Larsen, P.R., Mandel, S.J., Middleton, W.D., Reading, C.C. (2005). Management of thyroid nodules detected

at US: Society of radiologists in ultrasound consensus conference statement. *Radiology*, 237(3), 794–800.

- Friedrich, D., Jin, C., Zhang, Y., Demin, C., Yuan, L., Berynsky, L., Biesterfeld, S., Aach, T., & Böcking, A. (2012). Identification of prostate cancer cell nuclei for DNA-grading of malignancy. In: Tolxdorff, T.M., Deserno, H., & Meinzer H.P. (Hrsg.), *Bildverarbeitung für die Medizin*, Springer Berlin Heidelberg, 334–339.
- Friedrich, D. (2016). Effective improvement of cancer diagnostics and prognostics by computer-assisted cell image analysis. *Math-Nat. Diss.*, RWTH Aachen.
- Führer, D., & Schmid, K. W. (2010). Benigner Schilddrüsenknoten oder Schilddrüsenmalignom? *Der Internist*, 51(5), 611–619.
- Führer D., Bockisch, A., & Schmid, K. W. (2012). Euthyroid goiter with and without nodules-diagnosis and treatment. *Deutsches Ärzteblatt International*, 109(29–30), 506–515.
- Führer, D., Schmid, K.W., & Dralle, H. (2019). Anaplastisches Schilddrüsenkarzinom - Update 2019. *Der Onkologe*, 25(7), 569–572.
- Georgopoulos, N.A., Sykiotis, G.P., Sgourou, A., Papachatzopoulou, A., Markou, K.B., Kyriazopoulou, V., Papavassiliou, A.G., & Vagenakis, A.G. (2003). Autonomously functioning thyroid nodules in a former iodine-deficient area commonly harbor gain-of-function mutations in the thyrotropin signaling pathway. *European Journal of Endocrinology*, 149(4), 287–292.
- Gharib, H., Goellner, J. R., & Johnson, D. A. (1993). Fine-needle aspiration cytology of the thyroid. A 12-year experience with 11,000 biopsies. *Clinics in Laboratory Medicine*, 13(3), 699–709.
- Gharib, H., Papini, E., Garber, J.R., Duick, D.S., Harrell, R.M., Hegedüs, L., Paschke, R., Valcavi, R., Vitti, P., & AACE/ACE/AME Task Force on Thyroid Nodules. (2016). American association of clinical endocrinologists, American college of endocrinology and associazione medici endocrinologi medical guidelines for clinical practice for the the diagnosis and management of thyroid nodules, 2016 UPDATE. *Endocrine Practice*, 22(5), 622–639.
- Gharib, H., Papini, E., Paschke, R., Duick, D. S., Valcavi, R., Hegedüs, L., Vitti, P., & AACE/AME/ETA Task Force on Thyroid Nodules. (2010). American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules. *Endocrine Practice*, 16 Suppl 1, 1–43.
- 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. Analytical Cellular Pathology*, 17(4), 201–208.
- Giuffrida, D., & Gharib, H. (1995). Controversies in the management of cold, hot, and occult thyroid nodules. *The American Journal of Medicine*, 99(6), 642–650.
- Goldstein, R.E., Netterville, J. L., Burkey, B., & Johnson, J. E. (2002). Implications of follicular neoplasms, atypia, and lesions suspicious for malignancy diagnosed by

- fine-needle aspiration of thyroid nodules. *Annals of Surgery*, 235(5), 656–662; discussion 662-664.
- Grant, C. S., Hay, I. D., Ryan, J. J., Bergstralh, E. J., Rainwater, L. M., & Goellner, J. R. (1990). Diagnostic and prognostic utility of flow cytometric DNA measurements in follicular thyroid tumors. *World Journal of Surgery*, 14(3), 283–289; discussion 289-290.
- Graumann. (1953). Zur Standardisierung des Schiff'schen Reagenz. *Zeitschrift für Wissenschaftliche Mikroskopie* (61), 225-226.
- Grote, H.-J., Schmiemann, V., Sarbia, M., Böcking, A. (2003). DNA extraction from bronchial aspirates for molecular cytology: which method to take? *Analytical and Cellular Pathology* 25 (2), 83-88.
- Grouven, U., Bender, R., Ziegler, A., & Lange, S. (2007). Der Kappa-Koeffizient. *Deutsche Medizinische Wochenschrift*, 132(1), 65-68.
- Grußendorf, M. (2015). Schilddrüse 2011: 20. Konferenz über die menschliche Schilddrüse, Heidelberg. Therapie der diffusen und nodösen Struma im Wandel der Zeiten. 1. Auflage, Lehmanns Media, S. 37.
- Guerrero, M.A., Suh, I., Vriens, M.R., Shen, W.T., Gosnell, J., Kebebew, E., Duh, Q.-Y., & Clark, O. H. (2010). Age and tumor size predicts lymph node involvement in Hürthle cell carcinoma. *Journal of Cancer*, 2 (1), 23–26.
- Haider, A.S., Rakha, E. A., Dunkley, C., & Zaitoun, A. M. (2011). The impact of using defined criteria for adequacy of fine needle aspiration cytology of the thyroid in routine practice. *Diagnostic Cytopathology*, 39(2), 81–86.
- Haroske, G., Baak, J.P., 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. *Analytical Cellular Pathology*, 23(2), 89–95.
- 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. *Analytical Cellular Pathology*, 17(4), 189–200.
- Haugen, B.R., Alexander, E.K., Bible, K. C., Doherty, G.M., Mandel, S. J., Nikiforov, Y.E., Pacini, F., Randolph, G. W., Sawka, A. M., Schlumberger, M., Schuff, K. G., Sherman, S. I., Sosa, J. A., Steward, D. L., Tuttle, R. M., & Wartofsky, L. (2016). 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The american thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid*, 26(1), 1–133.
- Hegedüs, L., Bonnema S.J., & Bennedbaek, F. (2003). Management of simple nodular goiter: Current status and future perspectives. *Endocrine Reviews*, 24 (1), 102-132.
- Heide, W. (1982). Die Polyploidie der Zellkerne in Leber, Schilddrüse und Samenblasenepithel des Menschen während der 9. und 10. Lebensdekade und sie beeinflussende Faktoren, Med. Diss., Heinrich-Heine-Universität Düsseldorf.

- Hilgers R.-D. (2007). Einführung in die Medizinische Statistik. 2. Auflage, Springer Berlin, S. 34ff.
- Hinze, R., Gimm, O., Brauckhoff, M., Schneyer, U., Dralle, H., & Holzhausen, H. J. (2001). "Physiologische" und "neoplastische" C-Zell-Hyperplasien der Schilddrüse Morphologisch und biologisch distinkte Entitäten? *Der Pathologe*, 22(4), 259–265.
- Horii, A., Yoshida, J., Sakai, M., Okamoto, S., Honjo, Y., Mitani, K., Hattori, K., & Kubo, T. (1999). Ki-67 positive fractions in benign and malignant thyroid tumours: Application of flow cytometry. *Acta Oto-Laryngologica*, 119(5), 617–620.
- Hörmann, R. (2005). Schilddrüsenerkrankungen: Leitfaden für Praxis und Klinik. ABW Wissenschaftsverlag, S. 24–34.
- İnci, M. F., Özkan, F., Yüksel, M., Şalk, İ., & Şahin, M. (2013). The effects of sonographic and demographic features and needle size on obtaining adequate cytological material in sonography-guided fine-needle aspiration biopsy of thyroid nodules. *Endocrine*, 43(2), 424–429.
- Karger, S., Tötsch, M., Krause, K., & Führer, D. (2008). Der Schilddrüsenknoten - Differentialdiagnostik und Therapiekonzepte. *Medizinische Klinik*, 103(8), 575–590.
- Karges, W. (2010). Kalzitininbestimmung zur Frühdiagnose des medullären Schilddrüsenkarzinoms. *Der Chirurg*, 81(7), 620–626.
- Karges, W., Dralle, H., Raue, F., Mann, K., Reiners, C., Grussendorf, M., Hüfner, M., Niederle, B., Brabant, G., & German Society for Endocrinology (DGE) - Thyroid Section. (2004). Calcitonin measurement to detect medullary thyroid carcinoma in nodular goiter: German evidence-based consensus recommendation. *Experimental and Clinical Endocrinology & Diabetes*, 112(1), 52–58.
- Kashyap, V., Kaushik, N., Bhambhani, S., Das, D. K., & Luthra, U. K. (1992). Supportive role of image analysis and DNA ploidy pattern in the diagnosis of thyroid tumors. *Diagnostic Cytopathology*, 8(3), 228–230.
- Kindermann, D., & Hilgers, C. H. (1994). Glare-correction in DNA image cytometry. *Analytical Cellular Pathology*, 6(2), 165–180.
- Kloos, R. T. (2010). Thyroid cancer recurrence in patients clinically free of disease with undetectable or very low serum thyroglobulin values. *The Journal of Clinical Endocrinology and Metabolism*, 95(12), 5241–5248.
- Koss, L. G., & Melamed, M. R. (2005). Koss' Diagnostic cytology and its histopathologic bases. 5. Auflage, Verlag Lippincott, Williams & Wilkins, Philadelphia, S. 922.
- Krams, M., & Frahm, S. O. (2019). Kurzlehrbuch Pathologie. 3. Auflage, Georg Thieme Verlag, Stuttgart
- Krohn, K., Führer, D., Bayer, Y., Eszlinger, M., Brauer, V., Neumann, S., & Paschke, R. (2005). Molecular pathogenesis of euthyroid and toxic multinodular goiter. *Endocrine Reviews*, 26(4), 504–524.

- Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, & Jung HK. (2011). Thyroid imaging reporting and data system for US features of nodules: A step in establishing better stratification of cancer risk. *Radiology*, 260(3), 892-899.
- Labourier, E., Shifrin, A., Busseniers, A. E., Lupo, M. A., Manganelli, M. L., Andruss, B., Wylie, D., & Beaudenon-Huibregtse, S. (2015). Molecular testing for miRNA, mRNA, and DNA on fine-needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. *The Journal of Clinical Endocrinology and Metabolism*, 100(7), 2743–2750.
- Lan, L., Luo, Y., Zhou, M., Huo, L., Chen, H., Zuo, Q., & Deng, W. (2020). Comparison of diagnostic accuracy of thyroid cancer with ultrasound-guided fine-needle aspiration and core-needle biopsy: A systematic review and meta-analysis. *Frontiers in endocrinology*, 11, 44.
- Langer, J. E., Baloch, Z. W., McGrath, C., Loevner, L. A., & Mandel, S. J. (2012). Thyroid nodule fine-needle aspiration. *Seminars in Ultrasound, CT, and MR*, 33(2), 158–165.
- Leenhardt, L., Bernier, M. O., Boin-Pineau, M. H., Conte Devolx, B., Maréchaud, R., Niccoli-Sire, P., Nocaudie, M., Orgiazzi, J., Schlumberger, M., Wémeau, J. L., Chérie-Challine, L., & De Vathaire, F. (2004). Advances in diagnostic practices affect thyroid cancer incidence in France. *European Journal of Endocrinology*, 150(2), 133–139.
- Lehnert, H. (2014). *Rationelle Diagnostik und Therapie in Endokrinologie, Diabetologie und Stoffwechsel*. 4. Auflage, Georg Thieme Verlag, S.129.
- Lewis, C. M., Chang, K.-P., Pitman, M., Faquin, W. C., & Randolph, G. W. (2009). Thyroid fine-needle aspiration biopsy: Variability in reporting. *Thyroid: Official Journal of the American Thyroid Association*, 19(7), 717–723.
- Li, G., Cottier, M., Sabido, O., Gentil-Perret, A., Lambert, C., Genin, C., & Tostain, J. (2005). Different DNA ploidy patterns for the differentiation of common subtypes of renal tumors. *Cellular Oncology*, 27(1), 51–56.
- Li, H., Robinson, K. A., Anton, B., Saldanha, I. J., & Ladenson, P. W. (2011). Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. *The Journal of Clinical Endocrinology and Metabolism*, 96(11), 1719-1726.
- LiVolsi, V. A., & Baloch, Z. W. (2009). The many faces of follicular variant of papillary thyroid carcinoma. *Research Gate*, 14(6), 214–218.
- Lloyd, R. V., Osamura, R. Y., Klöppel G, Rosai J, & editors. (2017). *WHO classification of tumours of endocrine organs. Kapitel 2: Tumours of the thyroid gland*. 4. Auflage, IARC, Lyon, 65-143
- Lörincz, B. B., Simon, C., Möckelmann, N., & Knecht, R. (2016). Neue Aspekte in der Therapie von Schilddrüsenkarzinomen. *HNO*, 64(10), 736–740.
- Lundgren, C. I., Zedenius, J., & Skoog, L. (2008). Fine-needle aspiration biopsy of benign thyroid nodules: An evidence-based review. *World Journal of Surgery*, 32(7), 1247–1252.

- Magnanti, B., Dorak, M., Parker, L., Craft, A., James, P., & McNally, R. (2009). Geographical analysis of thyroid cancer in young people from northern England: Evidence for a sustained excess in females in Cumbria. *European Journal of Cancer*, 45(9), 1624–1629.
- Mahfouz, S. M., El-Sharkawy, S. L., Sharaf, W. M., El-din Hussein, H., & El-Nemr, R. S. (2012). Image cytometry of fine needle aspiration of thyroid epithelial lesions. *Applied Immunohistochemistry & Molecular Morphology*, 20(1), 25–30.
- Mann, K. (2002). Diagnostik und Therapie differenzierter Schilddrüsenkarzinome. *Der Internist*, 43(2), 174–185.
- Melchert, H-U., Görsch, B. & Thierfelder, W. (2002). Schilddrüsenhormone und Schilddrüsenmedikamente bei Probanden in den Nationalen Gesundheitssurveys. Berlin: Robert-Koch-Institut.
- Melchert, H-U., Görsch, B. & Thierfelder, W. (2005). Gebrauch von Schilddrüsen-therapeutika und Schilddrüsenhormon-Status. Eine vergleichende Auswertung von Verwendern und Nicht-Verwendern von Schilddrüsen-therapeutika unter den Probanden der drei Nationalen Surveys (NUS T0, T1, T2, sowie des Survey Ost). Berlin: Robert-Koch-Institut.
- Meng, Z., Shi, J., Zhu, C., Gu, J., & Zhou, C. (2013). Automated quantification of DNA aneuploidy by image cytometry as an adjunct for the cytologic diagnosis of malignant effusion. *Analytical Cellular Pathology*, 36(3–4), 107–115.
- Merchant, W. J., Thomas, S. M., Coppen, M. J., & Prentice, M. G. (1995). The role of thyroid fine needle aspiration (FNA) cytology in a district general hospital setting. *Cytopathology*, 6(6), 409–418.
- Montone, K. T., Baloch, Z. W., & LiVolsi, V. A. (2008). The thyroid Hürthle (oncocytic) cell and its associated pathologic conditions: A surgical pathology and cytopathology review. *Archives of Pathology & Laboratory Medicine*, 132(8), 1241–1250.
- Moore, M. D., Panjwani, S., Gray, K. D., Finnerty, B. M., Zarnegar, R., & Fahey, T. J. (2017). The role of molecular diagnostic testing in the management of thyroid nodules. *Expert Review of Molecular Diagnostics*, 17(6), 567–576.
- Mürbeth, S., Rousarova, M., Scherb, H., & Lengfelder, E. (2004). Thyroid cancer has increased in the adult populations of countries moderately affected by Chernobyl fallout. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 10(7), 300–306.
- Nadjari, B., Motherby, H., Pooschke, T., Pooschke, S., Gabbert, H.E., Simon, D., Röher, H.D., Feldkamp, J., Tharandt, L., Böcking, A. (1999). DNA aneuploidy as a specific marker of neoplastic cells in FNAB of the thyroid. *Analytical and Quantitative Cytology and Histology*, 21(6), 481–488.
- Nikiforov, Y. E., Seethala, R. R., Tallini, G., Baloch, Z. W., Basolo, F., Thompson, L. D. R., Barletta, J. A., Wenig, B. M., Al Ghuzlan, A., Kakudo, K., Giordano, T. J., Alves, V. A., Khanafshar, E., Asa, S. L., El-Naggar, A. K., Gooding, W. E., Hodak, S. P., Lloyd, R. V., Maytal, G., Ghossein, R. A. (2016). Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: A paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncology*, 2(8), 1023–1029.

- Nishino, M., & Krane, J. F. (2020). Role of ancillary Techniques in Thyroid Cytology Specimens. *Acta Cytologica*, 64(1–2), 40–51.
- Pacini, F., Schlumberger, M., Dralle, H., Elisei, R., Smit, J. W. A., Wiersinga, W., & European Thyroid Cancer Taskforce. (2006). European consensus for the management of patients with differentiated thyroid carcinoma of the follicular epithelium. *European Journal of Endocrinology*, 154(6), 787–803.
- Parfenova, E., Liu, K. Y. P., Harrison, A., MacAulay, C., Guillaud, M., & Poh, C. F. (2021). An improved algorithm using a Health Canada-approved DNA-image cytometry system for non-invasive screening of high-grade oral lesions. *Journal of Oral Pathology & Medicine*, 50(5), 502–509.
- Paschke, R., Cantara, S., Crescenzi, A., Jarzab, B., Musholt, T. J., & Sobrinho Simoes, M. (2017). European Thyroid Association guidelines regarding thyroid nodule molecular fine-needle aspiration cytology diagnostics. *European Thyroid Journal* 6, 115–129
- Paschke, R., Hegedüs, L., Alexander, E., Valcavi, R., Papini, E., & Gharib, H. (2011). Thyroid nodule guidelines: Agreement, disagreement and need for future research. *Nature Reviews Endocrinology*, 7(6), 354–361.
- Paschke, R., Schmid, K. W., & Gärtner, R. (2010). Epidemiologie, Pathophysiologie, Leitliniengerechte Diagnostik und Therapie des Schilddrüsenknotens. *Medizinische Klinik*, 105, 80-87.
- Pasha, H. A., Dhanani, R., Mughal, A., Ahmed, K. S., & Suhail, A. (2020). Malignancy rate in thyroid nodules with atypia or follicular lesion of undetermined significance. *International Archives of Otorhinolaryngology*, 24(2), e221–e226.
- Paunovic, I., Isic, T., Havelka, M., Tatic, S., Cvejic, D., & Savin, S. (2012). Combined immunohistochemistry for thyroid peroxidase, galectin-3, CK19 and HBME-1 in differential diagnosis of thyroid tumors. *Acta Pathologica, Microbiologica et Immunologica Scandinavica*, 120(5), 368–379.
- Perros, P., Boelaert, K., Colley, S., Evans, C., Evans, R. M., Gerrard Ba, G., Gilbert, J., Harrison, B., Johnson, S. J., Giles, T. E., Moss, L., Lewington, V., Newbold, K., Taylor, J., Thakker, R. V., Watkinson, J., Williams, G. R., & British Thyroid Association. (2014). Guidelines for the management of thyroid cancer. *Clinical Endocrinology*, 81 (1), 1–122.
- Pfannenstiel, P. (1993). Jodmangelstruma Diagnose- Therapie- Prävention. *Deutsch Ärzteblatt* 90 (15), 715-720.
- Pimpl, W., Galvan, G., Kogelnik, H. D., Manfreda, D., Niederle, B., Schlag, P. M., & Waclawiczek, H.-W. (2013). Struma maligna: Derzeitiger Stand in Diagnose und Therapie. Springer-Verlag, S. 49.
- Radespiel-Tröger, M., Batzler, W. U., Holleczeck, B., Luttmann, S., Pritzkeleit, R., Stabenow, R., Urbschat, I., Zeissig, S. R., & Meyer, M. (2014). Inzidenzzunahme des papillären Schilddrüsenkarzinoms in Deutschland. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 57(1), 84–92.
- Radl, Hudelist, & Kresnik. (2012). Doppler-Ultraschall in der Schilddrüsenabklärung: Eine Übersicht. *Journal für klinische Endokrinologie und Stoffwechsel*, 5(3), 7-10.

- Reiners, C., Dietlein, M., & Luster, M. (2008). Struma maligna - Schilddrüsenkarzinome. *Deutsche Medizinische Wochenschrift*, 133(43), 2215–2228.
- Reiners, C., Geling, M., Luster, M., Farahati, J., & Mäder, U. (2005). Epidemiologie des Schilddrüsenkarzinoms. *Der Onkologe*, 11(1), 11–19.
- Reiners, C., Wegscheider, K., Schicha, H., Theissen, P., Vaupel, R., Wrbitzky, R., & Schumm-Draeger, P.-M. (2004). Verbreitung von Schilddrüsenerkrankungen in der erwerbstätigen Bevölkerung Deutschlands: Ultraschall-Screening bei 96.278 Beschäftigten. *Thyroid*, 14: 926–932.
- Renshaw, A.A. (2011). Subclassification of atypical cells of undetermined significance in direct smears of fine-needle aspirations of the thyroid: Distinct patterns and associated risk of malignancy. *Cancer Cytopathology*, 119(5), 322–327.
- Rezaie, M. (2018). Erprobung der automatischen DNA-Bildzytometrie am Prostatakarzinom. *Med. dent. Diss*, Heinrich-Heine-Universität Düsseldorf.
- RKI (2019). Kurzzitierweise für: Robert-Koch-Institut (Hrsg.), & Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg.). *Krebs in Deutschland für 2015/2016*. 12. Ausgabe.
- Rosai, J. (2005). Handling of thyroid follicular patterned lesions. *Endocrine Pathology*, 16(4), 279–283.
- Rosato, L., Avenia, N., Bernante, P., De Palma, M., Gulino, G., Nasi, P. G., Pelizzo, M. R., & Pezzullo, L. (2004). Complications of thyroid surgery: Analysis of a multicentric study on 14,934 patients operated on in Italy over 5 years. *World Journal of Surgery*, 28(3), 271–276.
- Rosen, J. E., & Stone, M. D. (2006). Contemporary diagnostic approach to the thyroid nodule. *Journal of Surgical Oncology*, 94(8), 649–661.
- Rothmund, M. (2007). *Praxis der Viszeralchirurgie: Endokrine Chirurgie*. 3. Auflage, Springer Science & Business Media, S.52.
- Salajegheh, A., Petcu, E. B., Smith, R. A., & Lam, A. K.-Y. (2008). Follicular variant of papillary thyroid carcinoma: A diagnostic challenge for clinicians and pathologists. *Postgraduate Medical Journal*, 84(988), 78–82.
- Salmon, I., Gasperin, P., Rummelink, M., Rahier, I., Rocmans, P., Pasteels, J. L., Heimann, R., & Kiss, R. (1993). Ploidy level and proliferative activity measurements in a series of 407 thyroid tumors or other pathologic conditions. *Human Pathology*, 24(8), 912–920.
- Sandritter, W. (1964). Photometric cytodiagnosis of vaginal smears. *Journal of the National Cancer Institute*, 32, 1221–1228.
- Schäffer, R., Schmid, K. W., & Tötsch, M. (2012). Bethesda-Klassifikation der Feinnadelpunktion der Schilddrüse. *Der Pathologe*, 33(4), 324–330.
- Schmid, K. W. (2010). Pathogenese, Klassifikation und Histologie von Schilddrüsenkarzinomen. *Der Onkologe*, 16, 644–656.
- Schmid, K. W., & Farid, N. R. (2006). How to define follicular thyroid carcinoma? *Virchows Archiv*, 448(4), 385–393.

- Schmid, K. W., & Reiners C. (2011). Wann ist Feinnadelbiopsie der Schilddrüse am effektivsten? *Der Pathologe*, 32(2), 169–172.
- Schmid, W., & Tötsch, M. (2006). Dünnschichtzytologie der Schilddrüse. Handout Internationale Akademie für Pathologie (IAP), Symposium.
- Schmidt, D. (2016). Sonografie von Schilddrüsenknoten. *Der Nuklearmediziner*, 39(3), 166–171.
- Scott-Coombes, D., & Kinsman, R. (2007). The British Association of Endocrine Surgeons 2007: Second National Audit Report. Dendrite Clinical Systems, Oxford.
- Seagrove-Guffey, M. A., Hatic, H., Peng, H., Bates, K. C., & Odugbesan, A. O. (2018). Malignancy rate of atypia of undetermined significance/follicular lesion of undetermined significance in thyroid nodules undergoing FNA in a suburban endocrinology practice: A retrospective cohort analysis. *Cancer Cytopathology*, 126(10), 881–888.
- Seidel, A., & Sandritter, W. (1963). Cytophotometrische Messungen des DNS-Gehaltes eines pulmonalen Adenoms und einer malignen Lungenadenomatose. *Zeitschrift für Krebsforschung*, 65, 555–559.
- Shapiro S.S., & Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52 (3-4), 591–611.
- Shetty, S. K., Maher, M. M., Hahn, P. F., Halpern, E. F., & Aquino, S. L. (2006). Significance of incidental thyroid lesions detected on CT: Correlation among CT, sonography, and pathology. *American Journal of Roentgenology*, 187(5), 1349–1356.
- Siewert, J. R., Rothmund, M., & Schumpelick, V. (2012). *Praxis der Viszeralchirurgie: Endokrine Chirurgie*. 3. Auflage, Springer-Verlag, S.31.
- Sobrinho-Simões, M., Eloy, C., Magalhães, J., Lobo, C., & Amaro, T. (2011). Follicular thyroid carcinoma. *Modern Pathology*, 24(2), 10–18.
- Söderström, N. (1952). Puncture of goiters for aspiration biopsy. *Acta Medica Scandinavica*, 144(3), 237–244.
- Thamm, M., Ellert, U., Thierfelder, W., & Völzke, H. (2007). Jodversorgung in Deutschland - Ergebnisse des Jodmonitorings im Kinder- und Jugendgesundheitsurvey (KiGGS). *Bundesgesundheitsblatt* 50, 744-749.
- Thomas, C. (2006). *Histopathologie: Lehrbuch und Atlas zur Befunderhebung und Differenzialdiagnostik*. 14. Auflage, Schattauer Verlag, S.268.
- Tiedje, V., Ting, S., Dralle, H., Schmid, K. W., & Führer, D. (2015). Das medulläre Schilddrüsenkarzinom. *Der Internist*, 56(9), 1019–1031.
- Ting, S., Bockisch, A., Führer, D., Tötsch, M., & Schmid, K. W. (2012). Feinnadelbiopsie (FNB) der Schilddrüse. *Der Nuklearmediziner*, 35(01), 22–29.
- Ting, S., Synowacki, S., Bockisch A, Führer D, & Schmid K. W. (2015). Klinische Bedeutung der Schilddrüsenzytologie. *Der Pathologe*, 36(6), 543-552.
- Tosi, P., & Cottier, H. (1989). What's new in quantitative pathology? Past developments and outlook. *Pathology, Research and Practice*, 184(6), 652–655.

- Tötsch, M., Quadbeck, B., Görges, R., & Schmid, K. W. (2005). Präoperative Punktionszytologie beim Schilddrüsenkarzinom. *Der Onkologe*, 11(1), 40–49.
- Tribukait, B. (1993). Nuclear deoxyribonucleic acid determination in patients with prostate carcinomas: Clinical research and application. *European Urology*, 23 Suppl 2, 64–76.
- Tumorregister München (Hrsg.): ICD-10 C73: Schilddrüsenkarzinom. Survival. Online unter https://www.tumorregister-muenchen.de/facts/surv/sC73_G-ICD-10-C73-Schilddruesenkarzinom-Survival.pdf; abgerufen am 6.9.2021
- Völzke, H. (2012). Abschlussbericht. Monitoring der Prävalenz von Schilddrüsenerkrankungen in der Erwachsenenbevölkerung. SHIP-Trend 2008-2012.
- Völzke, H., Lüdemann, J., Robinson, D. M., Spieker, K. W., Schwahn, C., Kramer, A., John, U., & Meng, W. (2003). The prevalence of undiagnosed thyroid disorders in a previously iodine-deficient area. *Thyroid*, 13(8), 803–810.
- Wang, C., Friedman, L., Kennedy, G. C., Wang, H., Kebebew, E., Steward, D. L., Zeiger, M. A., Westra, W. H., Wang, Y., Khanafshar, E., Fellegara, G., Rosai, J., Livolsi, V., & Lanman, R. B. (2011). A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid*, 21(3), 243–251.
- Wienhold, R. (2013). The management of thyroid nodules: A retrospective analysis of health insurance data. *Deutsches Ärzteblatt International*, 110(49), 827–834.
- Wienke, J. R., Chong, W. K., Fielding, J. R., Zou, K. H., & Mittelstaedt, C. A. (2003). Sonographic features of benign thyroid nodules: Interobserver reliability and overlap with malignancy. *Journal of Ultrasound in Medicine*, 22(10), 1027–1031.
- World Health Organization, ICCIDD, & UNICEF. (2007). Assessment of iodine deficiency disorders and monitoring their elimination. 3. Auflage. Online unter https://apps.who.int/iris/bitstream/handle/10665/43781/9789241595827_eng.pdf;jsessionid=19B5D047D4B05FD7C34026CC13AEFD16?sequence=1, abgerufen am 6.9.2021
- Yang, J., Schnadig, V., Logrono, R., & Wasserman, P. G. (2007). Fine-needle aspiration of thyroid nodules: A study of 4703 patients with histologic and clinical correlations. *Cancer*, 111(5), 306–315.
- Zechmann C. M., & Haufe, S. E. (2012). Nuklearmedizinische und radiologische Diagnostik und Therapie der Schilddrüsenerkrankungen. Teil 1: Benigne Erkrankungen der Schilddrüse, *Der Radiologe*, 52, 661–678.

6 Abbildungsverzeichnis

Abbildung 1: Fertige, nach Feulgen gefärbte Präparate.....	21
Abbildung 2: Interaktives DNA-Bildzytometer MotiCyte DNA.....	23
Abbildung 3: Messmonitor des interaktiven Messsystems, hier mit einem Mund- schleimhautabstrich	24
Abbildung 4: Seite 1 des dreiseitigen Berichts einer manuellen DNA-zytometrischen Messung.....	25
Abbildung 5: Seite 2 des dreiseitigen Berichts einer manuellen DNA-zytometrischen Messung.....	26
Abbildung 6: Seite 3 des dreiseitigen Berichts einer manuellen DNA-zytometrischen Messung.....	27
Abbildung 7: Automatisches DNA-Bildzytometer (MotiCyte DNA Cytology Work- station auto)	28
Abbildung 8: Monitordarstellung der erzeugten Übersichtskarte des Objektträgers (rechts unten)	29
Abbildung 9: Monitordarstellung nach Setzen der Fokuspunkte (gelb) im zellhaltigen Messareal (grün)	30
Abbildung 10: Monitordarstellung des fertigen Scanvorgangs	31
Abbildung 11: Rohdaten eines gescannten Präparates.....	31
Abbildung 12: Präparat nach Umklassifizierung	32
Abbildung 13: Präparat nach Bearbeitung	32
Abbildung 14: Bildgalerie der Analysezellen nach Abschluss aller Bearbeitungen.....	33
Abbildung 15: Seite 1 des dreiseitigen Berichts einer automatischen DNA- zytometrischen Messung	34
Abbildung 16: Seite 2 des dreiseitigen Berichts einer automatischen DNA- zytometrischen Messung	35
Abbildung 17: Seite 3 des dreiseitigen Berichts einer automatischen DNA- zytometrischen Messung	36
Abbildung 18: Zwei repräsentative Beispiele von der Messung ausgeschlossener Präparate. Links: zu starke Abblassung nach mehrjähriger Lagerung; rechts: zu starke Hämorrhagie und Zellüberlappungen.	42
Abbildung 19: Alters- und Geschlechterverteilung im Patientenkollektiv (n=40)	43

Abbildung 20: Kreisdiagramm zur Verteilung der zytologischen Diagnosekategorien ...	44
Abbildung 21: Kreisdiagramm zur Darstellung der zytologischen Klartextdiagnosen	44
Abbildung 22: Kreisdiagramm zur Darstellung der histologischen diagnostischen Kategorien	45
Abbildung 23: Streudiagramm der Messwerte der minimalen C-Werte	48
Abbildung 24: Bland-Altman-Diagramm der Messwerte der minimalen C-Werte.....	48
Abbildung 25: Streudiagramm der Messwerte der maximalen C-Werte.....	50
Abbildung 26: Bland-Altman-Diagramm der Messwerte der maximalen C-Werte	50
Abbildung 27: Streudiagramm der Messwerte für den 5cEE (39 Messwertpaare, siehe Text)	52
Abbildung 28: Bland-Altman-Diagramm für den 5cEE (39 Messwertpaare, siehe Text)	52
Abbildung 29: Streudiagramm der Messwerte für den 5cER	54
Abbildung 30: Bland-Altman-Diagramm der Messwerte für den 5cER	54
Abbildung 31: Häufigkeitsverteilung der 9cEE im Vergleich manueller und auto- matischer Messungen	56
Abbildung 32: Streudiagramm der Messwerte für den 9cEE	57
Abbildung 33: Bland-Altman-Diagramm der Messwerte für den 9cEE.....	58
Abbildung 34: Streudiagramm der Messwerte für den 9cER	59
Abbildung 35: Bland-Altman-Diagramm der Messwerte für den 9cER	59
Abbildung 36: Streudiagramm der Messwerte der DNA-Stammlinie.....	61
Abbildung 37: Bland-Altman-Diagramm der Messwerte der DNA-Stammlinien	61
Abbildung 38: Histogramme der DNA-bildzytometrischen Messungen zu Fall 10. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.....	82
Abbildung 39: Histogramme der DNA-bildzytometrischen Messungen zu Fall 23. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.....	83
Abbildung 40: Histogramme der DNA-bildzytometrischen Messungen zu Fall 7. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.....	83
Abbildung 41: Histogramme der DNA-bildzytometrischen Messungen zu Fall 14. Links: manuelle Messung, rechts: automatische Messung. Jeweils Reproduktion der Original-Histogramme.....	84

7 Tabellenverzeichnis

Tabelle 1: Ablauf der Feulgen-Färbung.....	20
Tabelle 2: Prinzip der Vier-Felder-Tafel	38
Tabelle 3: Anteile der einzelnen zytologischen diagnostischen Kategorien an den untersuchten 40 Präparaten	43
Tabelle 4: Zytologische Klartextdiagnosen	44
Tabelle 5: Anteile der einzelnen histologischen diagnostischen Kategorien der 40 Fälle.....	45
Tabelle 6: Vierfeldertafel zur diagnostischen Treffsicherheit der Zytologie aller 40 Fälle.....	46
Tabelle 7: Statistische Kennwerte für die Zahl der gemessenen Analysezellen und Referenzzellen.....	46
Tabelle 8: Statistische Kennwerte für die mittlere Referenzzell-IOD der manuellen und der automatischen Messungen; die Angaben beziehen sich auf nur 39 statt auf 40 Fallvergleiche, da ein Wert bei der manuellen Messung irrtümlich nicht abgespeichert worden war.....	47
Tabelle 9: Statistische Kennwerte für den minimalen C-Wert der manuellen und der automatischen Messungen	47
Tabelle 10: Statistische Kennwerte für den maximalen C-Wert der manuellen und der automatischen Messungen	49
Tabelle 11: Statistische Kennwerte für den 5cEE der manuellen und der automati- schen Messungen	51
Tabelle 12: Statistische Kennwerte für den 5cER der manuellen und der automati- schen Messungen	53
Tabelle 13: Statistische Kennwerte für den 9cEE der manuellen und der automati- schen Messungen	55
Tabelle 14: Übereinstimmungsmatrix für die Werte des 9cEE bei manueller bzw. automatischer Messung; Chi ² -Test: $p < 0,0001$	56
Tabelle 15: Einteilung der Fälle mit 9cEE nach ihrer Häufigkeit bei beiden Messun- gen, n=40 (alle Fälle) bzw. n = 13 (Fälle mit wenigstens in einer der beiden Messungen 9cEE > 0)	57

Tabelle 16: Statistische Kennwerte für den 9cER der manuellen und der automatischen Messungen	58
Tabelle 17: Statistische Kennwerte für die DNA-Stammlinie der manuellen und der automatischen Messungen	60
Tabelle 18: Übereinstimmungsmatrix für die DNA-Stammlinien-Auswertung bei manueller bzw. automatischer Messung, Chi ² -Test: p = 0,0002	62
Tabelle 19: Übereinstimmungsmatrix für die Auswertung des 9cEE nach den Vorgaben der Einzelzellaneuploidie bei manueller bzw. automatischer Messung, Chi ² -Test: p = 0,0001	63
Tabelle 20: Übereinstimmungsmatrix für die kombinierte Auswertung von DNA-Aneuploidie bei manueller bzw. automatischer Messung, Chi ² -Test: p = 0,0052	63
Tabelle 21: Datenverteilung nach Lage der DNA-Stammlinie, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv	64
Tabelle 22: Treffsicherheit nach Lage der DNA-Stammlinie, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv	64
Tabelle 23: Datenverteilung nach Lage der DNA-Stammlinie, bezogen auf das Ergebnis der histologischen Untersuchung	65
Tabelle 24: Treffsicherheit nach Lage der DNA-Stammlinie, bezogen auf das Ergebnis der histologischen Untersuchung ("kein Tumor" gegen "Tumor")	65
Tabelle 25: Treffsicherheit nach Lage der DNA-Stammlinie, bezogen auf das Ergebnis der histologischen Untersuchung ("gutartig" gegen "bösartig")	65
Tabelle 26: Datenverteilung nach Zahl der 9cEE, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv	66
Tabelle 27: Treffsicherheit nach Zahl der 9cEE, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv	66
Tabelle 28: Datenverteilung nach Zahl der 9cEE, bezogen auf das Ergebnis der histologischen Untersuchung	67
Tabelle 29: Treffsicherheit nach Zahl der 9cEE, bezogen auf das Ergebnis der histologischen Untersuchung ("kein Tumor" gegen "Tumor")	67
Tabelle 30: Treffsicherheit nach Zahl der 9cEE, bezogen auf das Ergebnis der histologischen Untersuchung ("gutartig" gegen "bösartig")	67

Tabelle 31: Datenverteilung der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der zytologischen Untersuchung.....	68
Tabelle 32: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringen- der Verdacht, pos. = positiv	68
Tabelle 33: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der histologischen Untersuchung	69
Tabelle 34: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der histologischen Untersuchung ("kein Tumor" gegen "Tumor").....	69
Tabelle 35: Treffsicherheit der DNA-Bildzytometrie bei kombinierter Auswertung, bezogen auf das Ergebnis der histologischen Untersuchung ("gutartig" gegen "böartig").....	69
Tabelle 36: Datenverteilung der DNA-Stammlinien und der 9cEE-Werte für die manuelle und die automatische Messung bei den 14 interpretatorisch diskrepanten Fällen. Werte, die die Kriterien für DNA-Aneuploidie er- füllen, sind fettgedruckt wiedergegeben. Im Endergebnis der kombinier- ten Auswertung diskrepante Fälle sind rot dargestellt.	70
Tabelle 37: Treffsicherheit der DNA-Bildzytometrie nach Lage der DNA-Stamm- linie, des 9cEE und der kombinierten Auswertung, bezogen auf das zytologische Ergebnis. zw. = zweifelhaft, dV = dringender Verdacht, pos. = positiv	87
Tabelle 38: Treffsicherheit der DNA-Bildzytometrie nach Lage der DNA-Stamm- linie, des 9cEE und der kombinierten Auswertung, bezogen auf das histologische Ergebnis	88

Danksagung

Mein ganz besonderer Dank gilt Herrn Prof. Dr. med. Stefan Biesterfeld, ehemals Leiter des Schwerpunktes Cytopathologie der Heinrich-Heine-Universität Düsseldorf, für die Überlassung des Themas sowie für seine Unterstützung und Betreuung - auch später von Koblenz aus. Durch seine professionelle und engagierte Unterstützung konnte diese Dissertation überhaupt vollendet werden.

Ich möchte insbesondere meinem Betreuer Dr. Martin Schramm dafür danken, dass er mich bei dieser Arbeit betreut und unterstützt hat.

Außerdem danke ich allen Mitarbeiterinnen und Mitarbeitern des Schwerpunktes Cytopathologie der Heinrich-Heine-Universität Düsseldorf für die freundliche und kompetente Hilfe bei allen organisatorischen und technischen Angelegenheiten.

Mein größter Dank gilt meinen Eltern und Schwestern Yasemin, Semra und Büsra, die mich nicht nur bei dieser Arbeit, sondern in allen Lebenslagen unterstützt haben. Ohne eure Motivation und euren Glauben an mich hätte ich es nie geschafft.

Nicht zuletzt danke ich meinem Mann Alperen für die Hilfe besonders bei medizinischen Fragen und dass er sowohl in Zeiten der Dissertation als auch darüber hinaus immer für mich da ist und mich zum Lachen bringt.

Mein speziellster Dank gilt dem kleinen Mann in meinem Bauch, der mir während der anstrengenden Phase der endgültigen Fertigstellung dieser Dissertation die größte seelische Unterstützung war. Wir freuen uns sehr auf dich.

