

Aus dem
Institut für Pathologie und Neuropathologie
der Universität Tübingen
Abteilung Allgemeine und Molekulare Pathologie und
Pathologische Anatomie

**Multiparametrische Analyse des invasiven Mammakarzinoms
ohne speziellen Typ (NST) mit Hilfe von visueller
Auswertung und Bildanalyse**

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von
Schotte, Tilman

2025

| | |
|----------------------|--------------------------------|
| Dekan: | Professor Dr. B. Pichler |
| 1. Berichterstatter: | Professor Dr. F. Fend |
| 2. Berichterstatter: | Professorin Dr. E.-M. Grischke |
| Tag der Disputation: | 11.07.2025 |

Ich widme diese Arbeit
meiner Mutter und Sina

Inhaltsverzeichnis

| | | |
|-------|--|----|
| 1. | Einleitung | 13 |
| 1.1 | Das Mammakarzinom | 13 |
| 1.1.1 | Epidemiologie und Klinik, allgemeine Aspekte | 13 |
| 1.1.2 | Invasives Mammakarzinom ohne speziellen Typ (no special type, NST) | 14 |
| 1.1.3 | Tumorklassifizierung | 14 |
| 1.1.4 | Bedeutung von molekularen Analysen | 19 |
| 1.1.5 | Der Proliferationsmarker Ki67 | 19 |
| 1.1.6 | Der Mitoseindex | 22 |
| 1.1.7 | Der Mitosemarker PhH3..... | 23 |
| 1.1.8 | Bisherige multiparametrische Ansätze zur Verbesserung des Gradings vom invasiven Mammakarzinom NST | 25 |
| 1.2 | Digitale Analyse (DA) und künstliche Intelligenz (KI)..... | 26 |
| 1.2.1 | QuPath..... | 28 |
| 1.2.2 | Digitale Analyse von Ki67 mit Hilfe von QuPath..... | 29 |
| 1.2.3 | Digitale Analyse von PhH3..... | 30 |
| 1.3 | Fragestellung und Ziele..... | 32 |
| 2 | Material und Methoden | 34 |
| 2.1 | Die Kohorten | 34 |
| 2.1.1 | Entdeckungskohorte | 35 |
| 2.1.2 | Validierungskohorte..... | 35 |
| 2.1.3 | Follow-Up Daten | 37 |
| 2.2 | Angewendete Färbeverfahren..... | 39 |
| 2.2.1 | HE-Färbung | 39 |
| 2.2.2 | Immunhistochemische Färbung | 41 |
| 2.2.3 | Verwendete Geräte und Verbrauchsmaterialien..... | 45 |
| 2.3 | Visuelle Auswertung..... | 46 |
| 2.3.1 | Re-Grading anhand der HE-Schnitte | 46 |
| 2.3.2 | Visuelle Auswertung des Proliferationsindex (Ki67) | 47 |
| 2.3.3 | Visuelle Auswertung des Mitoseindex (PhH3)..... | 52 |
| 2.4 | Digitale Auswertung | 57 |
| 2.4.1 | Digitale Auswertung des Proliferationsindex (Ki67)..... | 57 |
| 2.4.2 | Digitale Auswertung des Mitoseindex (PhH3) | 60 |

| | | |
|-------|---|-----|
| 2.5 | Statistische Auswertung | 62 |
| 3 | Ergebnisse | 64 |
| 3.1 | Entdeckungskohorte | 64 |
| 3.1.1 | Re-Grading | 64 |
| 3.1.2 | Visuelle und digitale Auswertung des Ki67-Proliferationsindex ... | 65 |
| 3.1.3 | Visuelle und digitale Auswertung des PhH3 Mitoseindex..... | 70 |
| 3.1.4 | Clusteranalyse | 72 |
| 3.2 | Validierungskohorte..... | 75 |
| 3.2.1 | Re-Grading | 75 |
| 3.2.2 | Visuelle und digitale Auswertung des Ki67-Proliferationsindex ... | 76 |
| 3.2.3 | Visuelle und digitale Auswertung des PhH3 Mitoseindex..... | 86 |
| 3.2.4 | Clusteranalyse | 91 |
| 3.2.5 | Korrelation der Clusteranalyse zum klinischen Follow-Up | 93 |
| 4 | Diskussion | 97 |
| 4.1 | Vergleich verschiedener Methoden zur Auswertung von Ki67 und PhH3 | 98 |
| 4.1.1 | Visuelle Auswertung des Proliferationsindex (Ki67) - Schätzung vs. IKWG-Protokoll..... | 98 |
| 4.1.2 | Visuelle Mitosezählung (PhH3 vs. HE)..... | 99 |
| 4.1.3 | Visuelle vs. digitale Auswertung | 101 |
| 4.2 | Unterschied zwischen der Auswertung des Proliferations- und Mitoseindex von Stanzbiopsien und Resektaten | 105 |
| 4.2.1 | Auswertung des Proliferationsindex..... | 105 |
| 4.2.2 | Mitosezählung..... | 106 |
| 4.3 | Einfluss der Clusteranalyse auf die Tumorklassifizierung und den klinischen Verlauf | 107 |
| 4.4 | Fazit | 110 |
| 5 | Zusammenfassung | 113 |
| 6 | Literaturverzeichnis..... | 115 |
| 7 | Erklärung zu Eigenanteil..... | 124 |
| 8 | Veröffentlichungen | 125 |
| 9 | Danksagung..... | 126 |

Abbildungs- und Tabellenverzeichnis

Abbildungen

| | |
|---|----|
| Abbildung 1: Schematische Darstellung der HER2-Testalgorithmen nach S3-Leitlinie Mammakarzinom..... | 18 |
| Abbildung 2: Übersicht der fünf Fachbereiche welche QuPath am häufigsten nutzen..... | 28 |
| Abbildung 3: Darstellung der Funktionsweisen des OptiView DAB Detection Kit und des VENTANA BenchMark ULTRA..... | 43 |
| Abbildung 4: Darstellung der einzelnen Mitosefiguren im HE-Schnitt zugeordnet zu den entsprechenden Teilungsschritten | 46 |
| Abbildung 5: Formel zur Berechnung des gewichteten Ki67-Scores gemäß dem Protokoll der IKWG..... | 47 |
| Abbildung 6: Analyse der Ki67-positiven Zellen in Bezug auf ihre Farbintensität gemäß den Empfehlungen der IKWG | 48 |
| Abbildung 7 Darstellung der Mib1-Färbung bei Fällen mit G1-, G2- und G3-Grading im Vergleich zur HE-Färbung derselben Regionen..... | 51 |
| Abbildung 8: Darstellung von Mitosefiguren und Färbeartefakten im PhH3-Schnitt | 53 |
| Abbildung 9 Darstellung der PhH3-Färbung bei Fällen mit G1-, G2- und G3-Grading im Vergleich zur HE-Färbung derselben Regionen..... | 56 |
| Abbildung 10: Zelldetektionstool der Software QuPath zur Auswertung de Ki67-Proliferationsindex..... | 59 |
| Abbildung 11: Zelldetektionstool der Software QuPath zur Auswertung der Mitoserate anhand von PhH3-gefärbten Schnitten..... | 61 |
| Abbildung 12: Sankey-Plot zur Darstellung der Veränderungen des Gradings nach den einzelnen Auswertungsmethoden (Entdeckungskohorte)..... | 64 |
| Abbildung 13: Boxplots zur Darstellung der Proliferationsindizes durch visuelle und digitale Auswertung (Entdeckungskohorte) | 66 |
| Abbildung 14: Gestapeltes Balkendiagramm mit der Aufteilung der visuellen und digitalen Methoden nach Gruppen unterschiedlicher Proliferationsraten in % (Entdeckungskohorte) | 67 |
| Abbildung 15: Darstellung der Korrelationen nach Pearson von digitalen und visuellen Methoden zur Auswertung des Ki67-Proliferationsindex (Entdeckungskohorte) | 68 |
| Abbildung 16: Säulendiagramm zum Vergleich der unterschiedlichen Auswertungsmethoden des Ki67 Proliferationsindex und den resultierenden Aufteilungen der Fälle zu den molekularen Subtypen Luminal A und B | 70 |
| Abbildung 17: Vergleich der Mitoseraten durch visuelle Auswertung von HE- und PhH3 gefärbten Schnitten sowie mithilfe der digitalen Methode (QuPath) (Entdeckungskohorte) | 71 |
| Abbildung 18: Scatter Plots mit den Korrelationen nach Pearson zwischen den einzelnen Mitoseauswertungen (Entdeckungskohorte) | 72 |
| Abbildung 19: Clusteranalyse der Entdeckungskohorte | 73 |

| | |
|---|----|
| Abbildung 20: Boxplots zur Darstellung der Verteilung der einzelnen Parameter innerhalb der Cluster (Entdeckungskohorte) | 74 |
| Abbildung 21: Sankey-Plot zur Darstellung der Veränderungen des Gradings nach den einzelnen Auswertungsmethoden (Validierungskohorte) | 75 |
| Abbildung 22: Vergleich der Proliferationsindizes zwischen visuellen und digitalen Methoden (Resektate, Validierungskohorte)..... | 77 |
| Abbildung 23: Gestapeltes Balkendiagramm mit der Aufteilung der visuellen und digitalen Methoden nach Gruppen unterschiedlicher Proliferationsraten in % (Resektate Validierungskohorte)..... | 78 |
| Abbildung 24: Scatter-Plots zur Darstellung der Korrelation nach Pearson der verwendeten Auswertungsmethoden des Ki67-Proliferationsindex (Resektate, Validierungskohorte)..... | 79 |
| Abbildung 25: Gestapeltes Balkendiagramm mit der Aufteilung der visuellen und digitalen Methoden nach Gruppen unterschiedlicher Proliferationsraten in % (Resektate Validierungskohorte)..... | 81 |
| Abbildung 26: Scatter-Plots zur Darstellung der Korrelation nach Pearson der verwendeten Auswertungsmethoden des Ki67-Proliferationsindex (Stanzbiopsien, Validierungskohorte) | 82 |
| Abbildung 27: Vergleich der Ki67-Proliferationsindizes zwischen den einzelnen Auswertungsmethoden an Stanzbiopsien und Resektaten (Validierungskohorte) | 84 |
| Abbildung 28: Scatter-Plots zur Darstellung der Korrelation nach Pearson der verwendeten Auswertungsmethoden des Ki67 Proliferationsindex zwischen Resektaten und Stanzbiopsien..... | 85 |
| Abbildung 29: Balkendiagramm zur Veranschaulichung der Aufteilung der molekularen Subtypen bei Resektaten und Stanzbiopsien unter Anwendung visueller und digitaler Auswertungsmethoden | 86 |
| Abbildung 30: Vergleich der Mitoseraten anhand von HE- und PhH3-Färbung | 87 |
| Abbildung 31: Scatter-Plot zur Darstellung der Mitoseraten anhand von HE- und PhH3-gefärbten Schnitten (Resektat)..... | 88 |
| Abbildung 32: Vergleich von digitaler und visueller Auswertung der Mitoseraten von Resektaten und Stanzbiopsien (PhH3)..... | 89 |
| Abbildung 33: Scatterplots zur Darstellung der Korrelation nach Pearson von digitaler und visueller PhH3-Auswertung bei Resektaten und Stanzbiopsien... | 89 |
| Abbildung 34: Vergleich der visuellen Auswertung der Mitoseraten von Resektaten und Stanzbiopsien (HE, PhH3)..... | 90 |
| Abbildung 35: Scatterplots zur Darstellung der Korrelation nach Pearson der visuellen Auswertungsmethoden für die Mitoseraten von Stanzbiopsien und Resektaten (HE, PhH3)..... | 91 |
| Abbildung 36: Clusteranalyse der Validierungskohorte | 92 |
| Abbildung 37: Boxplots zur Darstellung der Verteilung der einzelnen Parameter innerhalb der Cluster | 93 |
| Abbildung 38: Kaplan-Meyer-Diagramme die Grading (A) und multiparametrische Cluster (B) in Relation zum Rezidivfreien Überleben setzen | 94 |

Abbildung 39: Kaplan-Meyer-Diagramme, die Grading (A) und multiparametrische Cluster (B) in Relation zum Gesamtüberleben setzen..... 96

Tabellen

| | |
|---|----|
| Tabelle 1: Aufteilung der Mitosescores in Abhängigkeit zum GFD bei 40x Vergrößerung nach Ellis et al | 16 |
| Tabelle 2: Zusammensetzung des Patientenkollektivs der ersten Kohorte..... | 35 |
| Tabelle 3: Zusammensetzung des Patientenkollektivs der zweiten Kohorte..... | 37 |
| Tabelle 4: Protokoll zur automatisierten HE-Färbung mit dem Tissue-Tek Färbeautomaten | 39 |
| Tabelle 5: Verwendete Reagenzien zur Herstellung der Hämaluanvorratslösung nach Mayer..... | 40 |
| Tabelle 6: Verwendete Reagenzien zur Herstellung der Eosinvorratslösung (1% wässrig)..... | 40 |
| Tabelle 7: Zusätzliche Reagenzien und Lösungen, die für die Herstellung eines HE-Schnittes notwendig sind..... | 41 |
| Tabelle 8: Verwendete Antikörper für Ki67 und PhH3..... | 41 |
| Tabelle 9: Verwendete Reagenzien für die Immunhistochemischen Färbungen . | 44 |
| Tabelle 10: Zusammensetzung des OptiView Deteciton Kit laut Hersteller Roche | 45 |
| Tabelle 11: Verwendete Geräte für HE-und immunhistochemische Färbungen. | 45 |
| Tabelle 12: Verwendete Objektträger und Coverslips | 45 |
| Tabelle 13: Anzahl der Fälle mit schlechter oder fehlender Reaktion auf PhH3-Färbung..... | 53 |
| Tabelle 14: Re-Klassifikation nach Re-Evaluation und PhH3-Auswertung (Entdeckungskohorte) | 65 |
| Tabelle 15: Aufteilung der molekularen Subgruppen nach visueller und digitaler Analyse des Proliferationsindex (Ki67) (Entdeckungskohorte) | 69 |
| Tabelle 16: Re-Klassifikation nach Re-Evaluation und PhH3-Auswertung (Validierungskohorte)..... | 76 |
| Tabelle 17: Aufteilung der molekularen Subgruppen nach visueller und digitaler Analyse des Proliferationsindex (Ki67) am Resektat (Validierungskohorte) . | 80 |
| Tabelle 18: Aufteilung der molekularen Subgruppen nach visueller und digitaler Analyse des Proliferationsindex (Ki67) an den Stanzbiopsien (Validierungskohorte)..... | 84 |

Abkürzungen

| | |
|--------|---|
| - | Negativ |
| + | Positiv |
| AUC | Area Under the Curve |
| BET | Brust erhaltende Therapie |
| CCC | Comprehensive Cancer Center in Tübingen |
| DA | Digitale Analyse |
| DCIS | Duktales Carcinoma in Situ |
| EGF | Epidermale Wachstumsfaktoren |
| ER | Östrogenrezeptor |
| ErbB2 | ErbB2 receptor tyrosine kinase 2 |
| FFPE | Formalin fixiert und in Paraffin eingebettet |
| GFD | Gesichtsfelddurchmesser |
| GFF | Gesichtsfeldfläche |
| GRK | Graduiertenkolleg |
| HE | Hämatoxylin-Eosin-Färbung |
| HER2 | Human epidermal growth factor receptor 2 |
| HPF | Hochauflösende Gesichtsfelder/High-Power-Fields |
| HR | Hormonrezeptor |
| ICC | Intraclass correlation coefficient |
| IHC | Immunhistochemie |
| IKEaB | Institut für klinische Epidemiologie und angewandte Biometrie |
| IKWG | Internationale Ki67 Working Group |
| IQ | Interquartilabstand |
| ISH | In Situ Hybridisierung |
| KI | Künstliche Intelligenz |
| Ki67 | Kiel-67-Antigen |
| LCS | Liquid Cover Slip |
| M | Median |
| MAI | Mitotischer Aktivitätsindex |
| MIB1 | Made in Borstel 1-Antikörper |
| MTA | Medizinisch-technische Assistent*innen |
| MW | Mittelwert |
| NST | No special type |
| OB | Originalbefund |
| OS | Gesamtüberleben |
| PhH3 | Phosphohiston 3 |
| PR | Progesteronrezeptor |
| p-Wert | Probilitas-Wert |
| Q | Quartil |
| RFS | Rezidivfreies Überleben |

R-Wert
Ser
SPSS

TN
TNM-Klassifikation

WSI

Korrelationskoeffizient
Serin
Statistical Package for the Social
Sciences
Triple negativ
Tumour-Node-Metastasis-
Klassifikation
Whole Slide Image

1. Einleitung

1.1 Das Mammakarzinom

1.1.1 Epidemiologie und Klinik, allgemeine Aspekte

Das Mammakarzinom ist mit ca. 72.000 Fällen pro Jahr (Stand 2018) die am häufigsten auftretende maligne Tumorerkrankung bei Frauen in Deutschland. Betroffen ist dadurch in etwa jede achte Frau (Katalinic et al., 2020). Insgesamt macht es ungefähr 30% aller Krebsneuerkrankungen im Land aus (Dannehl and Hartkopf, 2022). Mit etwa 80% der Betroffenen manifestiert sich das Mammakarzinom am häufigsten bei Patientinnen über 50 Jahre (Łukasiewicz et al., 2021). Die Symptome können sich hierbei sehr unterschiedlich darstellen. In den meisten Fällen wird aber ein „Knoten“ in der Brust als erster Befund beschrieben. Interessanterweise suchen sich Patient:innen mit einem solchen Tastbefund im Durchschnitt 90 Tage früher Hilfe als Betroffene, die nicht brustspezifische Symptome aufweisen (Mamillenanomalien, Gewichtsverlust, Brust- oder Rückenschmerzen) (Koo et al., 2017).

Seit der Einführung des Mammographie-Screenings im Jahr 2009 für Frauen im Alter zwischen 50 und 69 Jahren im zweijährigen Turnus ist die Inzidenz von Mammakarzinomen hierzulande um 17-19% gesunken (Katalinic et al., 2020). Des Weiteren werden Tumoren immer häufiger in frühen Stadien entdeckt. Es wird dementsprechend vermutet, dass diese Präventionsmaßnahme einen positiven Einfluss auf die Mortalität haben könnte (Katalinic et al., 2020).

Trotzdem stellen Brustkrebserkrankungen einen erheblichen Kostenfaktor für das Gesundheitssystem dar. Die durchschnittlichen Behandlungskosten pro Patient:in belaufen sich in etwa auf 7.500 € (Schöffski and Emmert, 2009). Dies unterstreicht das fortwährende Interesse an der Optimierung von Therapiemaßnahmen.

In der Regel besteht die Standardtherapie aus einer operativen Entfernung des Tumors und gegebenenfalls betroffener Lymphknoten. Dabei kann eine brusterhaltende Operation (BET) oder eine Mastektomie durchgeführt werden. Abhängig von Größe, Lymphknotenbefund und molekularem Subtyp können

zusätzliche Maßnahmen wie Hormontherapie, Chemotherapie und Bestrahlung in Betracht gezogen werden (Moo et al., 2018).

1.1.2 Invasives Mammakarzinom ohne speziellen Typ (no special type, NST)

Das invasive Karzinom NST ist der häufigste Subtyp unter den Mammakarzinomen und macht etwa 50-80% aller Fälle aus (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021). Diese Tumorentität zeigt eine bemerkenswert heterogene klinische Erscheinung mit einem stark variierenden Anteil der Tumormasse und unterschiedlicher Tumorgröße. Die Differenzierung kann von duktalem Wachstum bis hin zu einer soliden Form reichen. Häufig weisen die Tumorzellen eine höhergradige Kernpleomorphie auf und in vielen Fällen können vermehrte Mitosefiguren erkannt werden (Makki, 2015).

Aufgrund der hohen Inzidenz wird das invasive Mammakarzinom NST seit vielen Jahren intensiv erforscht. Besonders die Anpassung und Optimierung der Therapie stehen im klinischen Alltag im Fokus. Mit dem Aufkommen der personalisierten Medizin erhalten auch Brustkrebspatient:innen zunehmend individuell abgestimmte Behandlungsansätze. Das Ziel ist eine möglichst effektive Therapie mit minimalen Nebenwirkungen, wobei Operationstechniken immer präziser gestaltet oder Bestrahlungszeiträume verkürzt werden (Goldhirsch et al., 2013).

Eine präzise und effektive Therapie baut jedoch auf einer genauen Klassifizierung des Tumors auf, die den Grundstein für eine optimierte Behandlungsstrategie legt.

1.1.3 Tumorklassifizierung

1.1.3.1 Tumour-Node-Metastasis-Klassifikation (TNM)

Die Klassifizierung von Mammakarzinomen erfolgt zunächst nach der TNM-Klassifikation, die auch für viele andere Karzinomformen angewendet wird. Dabei werden zunächst klinisch und später pathologisch die Tumorgröße (T), der regionäre Lymphknotenbefall (N) und das Vorliegen von Fernmetastasen (M)

beurteilt. Die Tumorgröße wird in T1 für Tumoren mit maximal 2 cm Durchmesser eingestuft, während T4 Tumoren jeder Größe umfasst, die sich direkt in Brustwand oder Haut ausdehnen. Das Fehlen von Metastasen in den regionären Lymphknoten wird als N0 bezeichnet. Wenn die ipsilateralen, infraklavikulären Lymphknoten mit oder ohne axilläre Lymphknoten befallen sind, wird dies im Befund mit N3 gekennzeichnet. Der Lymphknotenstatus wird durch klinische Untersuchung oder bildgebende Verfahren festgestellt. Abschließend wird beurteilt, ob Fernmetastasen vorhanden sind, die am häufigsten in Knochen, Lunge, Leber oder Gehirn auftreten können (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021, Jin and Mu, 2015).

1.1.3.2 Morphologisches Grading

Die Frage nach einer noch präziseren Einteilung von Mammakarzinomen beschäftigt Wissenschaftler bereits seit vielen Jahrzehnten. Basierend auf der Arbeit von Bloom und Richardson et al. (Bloom and Richardson, 1957) wurde 1991 die Klassifikation von Elston und Ellis et al. (Elston and Ellis, 1991) als Goldstandard etabliert. Diese Methode ermöglichte erstmals eine einheitliche Beschreibung der Differenzierung von Mammakarzinomen und hat in dieser Form bis heute Bestand.

Die Bewertung der Tumoren erfolgt anhand der Ausprägung der tubulären Formation, dem nukleären Pleomorphismus und der Mitosezahl. Jede Kategorie erhält dabei einen Score von 1-3, wodurch der höchstmögliche Gesamt-Score 9 beträgt. Ein Score von 9 zeigt an, dass der Tumor sich morphologisch am weitesten von dem Ursprungsgewebe entfernt hat.

Die tubuläre Formation wird in Prozent bewertet. Tumoren, die mehr als 75% Drüsenstrukturen aufweisen, erhalten den Score 1, während eine Ausprägung zwischen 10 und 75% mit dem Score 2 bewertet wird. Ein Score von 3 wird vergeben, wenn weniger als 10% dieser Formationen feststellbar sind.

Der zweite Parameter betrifft die Kernpleomorphien. Ein Score von 1 beschreibt kleine, reguläre, uniforme Zellen. Moderat variierende Größen werden mit Score 2 bewertet, während Score 3 deutliche Abweichungen im Erscheinungsbild der Zellen kennzeichnet.

Die Bestimmung der Mitosezahl erfolgt in der Regel über zehn hochauflösende Gesichtsfelder (High-Power-Fields, HPF), was einer Fläche von 1,25 bis 3,85 mm², abhängig vom Modell des Mikroskops, entspricht. Anhand der Anzahl der Mitosen pro 10 HPF wird der Mitotische-Aktivitätsindex (MAI) berechnet und in einen Score umgewandelt. Dabei ist der Gesichtsfelddurchmesser (GFD) des Mikroskops bei 40x Vergrößerung von entscheidender Bedeutung. Eine detaillierte Übersicht dazu kann aus Tabelle 1 entnommen werden:

| GFD (mm) | Zahl der Mitosen entsprechend | | |
|----------|-------------------------------|---------|---------|
| | Score 1 | Score 2 | Score 3 |
| 0,40 | ≤4 | 5-9 | ≥10 |
| 0,42 | ≤5 | 6-10 | ≥11 |
| 0,44 | ≤5 | 6-11 | ≥12 |
| 0,46 | ≤6 | 7-12 | ≥13 |
| 0,48 | ≤6 | 7-13 | ≥14 |
| 0,50 | ≤7 | 8-14 | ≥15 |
| 0,52 | ≤7 | 8-15 | ≥16 |
| 0,54 | ≤8 | 8-16 | ≥17 |
| 0,56 | ≤8 | 9-17 | ≥18 |
| 0,58 | ≤9 | 10-19 | ≥20 |
| 0,60 | ≤10 | 10-20 | ≥21 |
| 0,62 | ≤11 | 12-22 | ≥23 |
| 0,64 | ≤11 | 12-23 | ≥24 |

Tabelle 1: Aufteilung der Mitosescores in Abhängigkeit zum GFD bei 40x Vergrößerung nach Ellis et al (Ellis, 2016). Mit zunehmendem GFD sind mehr Mitosen erforderlich, um einen höheren Score zu erreichen

1.1.3.3 Molekulare Subtypen

Mit dem Aufkommen molekularer Analysemethoden boten sich völlig neue Möglichkeiten, Mammakarzinome noch präziser zu klassifizieren. Im Jahr 2000 konnten Perou et al. durch die systematische Untersuchung von Genexpressionsmustern vier Subtypen definieren: den Östrogenrezeptorpositiven (ER+)/Luminalen, den basal-like, den Erb-B2 Rezeptor Tyrosinkinase-2- (Erb-B2, später Human epidermal growth factor receptor 2 (HER2)) und den normalen Brust Typ (Perou et al., 2000).

Diese Klassifikation wurde später modifiziert, sodass wir heute die molekularen Subtypen wie folgt einteilen (Yersal and Barutca, 2014):

Die luminalen Tumoren werden anhand von Surrogatmarkern wie der Expression der Östrogen- und Progesteronrezeptoren (ER und PR) sowie des Proliferationsmarkers Ki67 (Ki67) in zwei Subgruppen eingeteilt. Luminal A Tumoren sind ER+ und $\geq 20\%$ PR+. Zusätzlich weisen sie eine Expression von Ki67 in $< 20\%$ der Tumorzellen auf. Bei Tumoren, bei denen PR $< 20\%$ und/oder Ki67 $\geq 20\%$ evaluiert wird, spricht man von einem Luminal B Tumor (Goldhirsch et al., 2013, Robertson et al., 2019). In einigen Studien werden Luminal B Tumoren, bei denen man eine Amplifikation des HER2-Gens feststellen kann auch als Luminal C (Tran and Bedard, 2011) oder HER2 enriched (Li et al., 2016) bezeichnet.

HER2 ist eine Tyrosinkinase aus der Familie der epidermalen Wachstumsfaktoren (EGF) (Mitri et al., 2012, Yersal and Barutca, 2014). Dieser Rezeptor wird mit einer schlechteren Prognose und einem aggressiveren Krankheitsverlauf von Brustkrebs in Verbindung gebracht (Mitri et al., 2012). Wenn ein Tumor hormonrezeptor-negativ (HR-) ist und gleichzeitig eine HER2-Positivität aufweist, gilt er als HER2+. Die visuelle Beurteilung der Rezeptor-Positivität erfolgt anhand klar definierter Algorithmen (siehe Abbildung 1). Wenn sich eine intermediäre oder schwache kontinuierliche Membrandfärbung zeigt, muss eine In Situ-Hybridisierung (ISH) durchgeführt werden, um die Amplifikation des HER2-Gens festzustellen (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021).

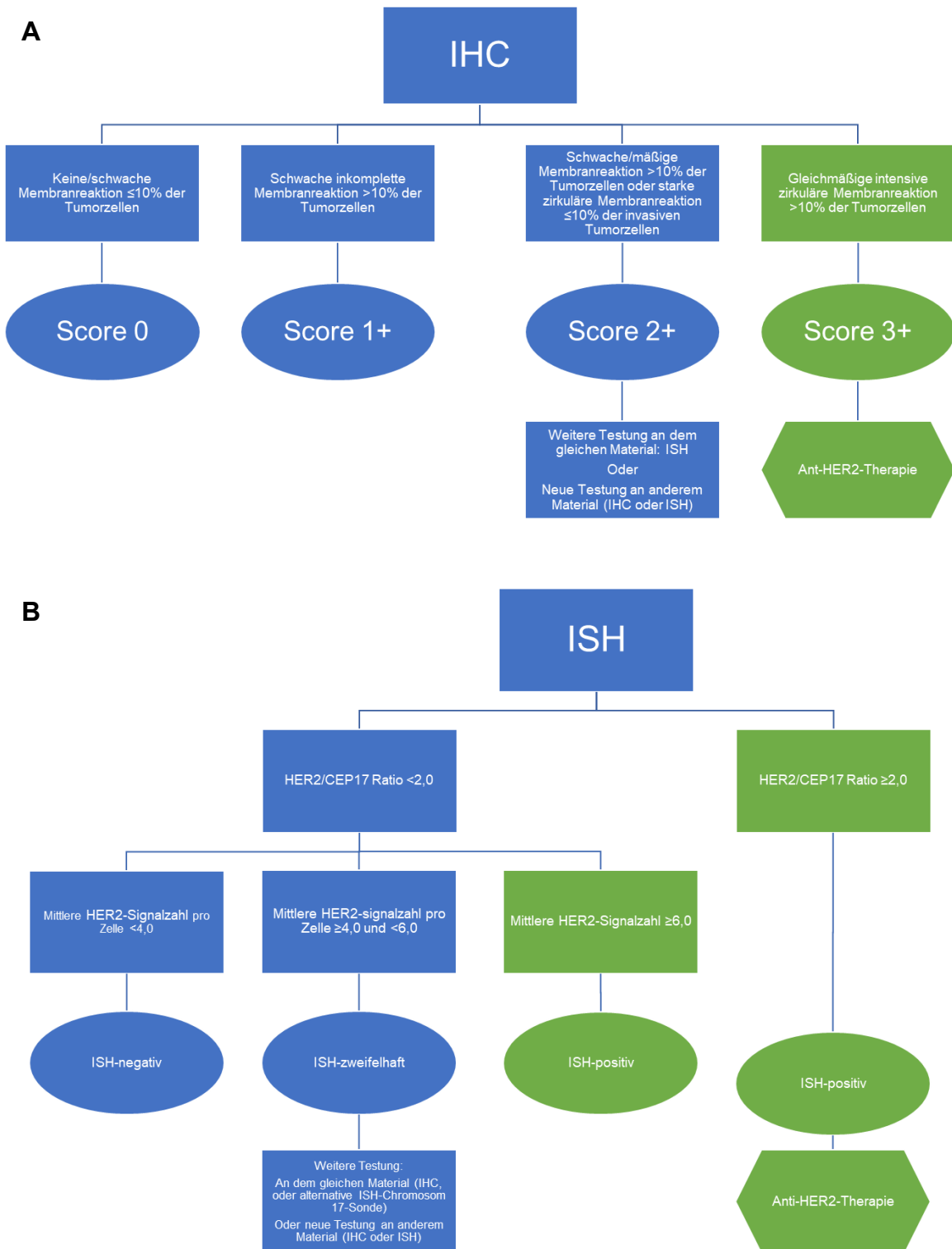


Abbildung 1: Schematische Darstellung der HER2-Testalgorithmen nach S3-Leitlinie Mammakarzinom. Zunächst wird der HER2-Status eines Präparats mittels Immunhistochemie (IHC) (A) bestimmt, bei zweifelhafter Membranfärbung erfolgt die Bestätigung mithilfe von ISH (B) (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021)

Die letzte Gruppe bilden die triple negativen (TN) Tumore. Sie zeichnen sich durch Nicht-Expression von HR und HER2 aus und enthalten auch die Gruppe der basal-like Karzinome. In der klinischen Praxis wird die erste Definition verwendet, da sie die IHC-Eigenschaften des Tumors beschreibt. Die Bezeichnung basal-like wird dann genutzt, wenn Fälle durch Genexpression analysiert wurden. Daher sollten diese Bezeichnungen nicht als Synonyme verwendet werden (Yersal and Barutca, 2014). Diese Tumor-Subgruppe gilt als besonders aggressiv und therapeutisch resistent (Aysola et al., 2013, Yersal and Barutca, 2014).

Die Häufigkeit der einzelnen Subtypen variiert je nach Studie: Luminal A (50-60%), Luminal B (15-20%) (darunter Luminal B HER2+ (ca. 6%)), HER2+ (15-20%) und basal-like bzw. TN (8-37%) (Yersal and Barutca, 2014, Tagliabue et al., 2021)

Neben dem Grading spielt die molekulare Klassifikation heute eine mindestens ebenbürtige Rolle, insbesondere im Zusammenhang mit der medikamentösen Therapie. Für HR+ Tumoren gilt die endokrine Systemtherapie als Grundlage der Behandlung (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021). Bei Hochrisiko-Typen mit einem hohen Ki67 Proliferationsindex >30% oder amplifiziertem HER2-Gen wird eine zusätzliche Chemotherapie in Betracht gezogen. Im Zusammenhang mit HER2+ Tumoren hat sich Trastuzumab, ein „humanisierter, monoklonaler, IgG1 Antikörper“ (Boekhout et al., 2011), als wirksame Therapieoption etabliert (Guarneri and Conte, 2009). Für TN-Tumoren sind die Therapieoptionen bisher begrenzt. Der aktuelle Konsens besagt, dass die Komplettremissionsrate dieser Entität durch Platinsalze erhöht wird (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021). In dieser Gruppe von Patient:innen werden auch immun-onkologische Therapiekonzepte umgesetzt.

1.1.4 Bedeutung von molekularen Analysen

1.1.5 Der Proliferationsmarker Ki67

Ki67 ist ein nukleäres Antigen, das während den aktiven Phasen des Zellzyklus in allen proliferierenden Zellen nachgewiesen werden kann (Gerdes et al., 1983,

de Azambuja et al., 2007). Es handelt sich um ein Nicht-Histon-Protein mit einem molekularen Gewicht von 395 KD (de Azambuja et al., 2007, Gerdes et al., 1991). Für die klinische Anwendung hat sich der Antikörper Made in Borstel 1 (MIB1) etabliert, der eine hohe Korrelation mit dem Ki67-Antigen aufweist (de Azambuja et al., 2007, Cattoretti et al., 1992).

Im Jahr 2007 veröffentlichte de Azambuja et al. eine Metaanalyse, in der sie feststellten, dass eine hohe Ki67-Positivität möglicherweise ein prognostischer Faktor für ein schlechteres Überleben und ein höheres Rezidivrisiko sein könnte (de Azambuja et al., 2007). Seitdem war dieser IHC-Marker immer wieder Gegenstand von Diskussionen. Kritiker:innen sehen vor allem Probleme in der fehlenden Standardisierung der Laborprozesse und der Interpretation der Befunde (Dowsett et al., 2011). Aus diesem Grund wurde 2009 die internationale Ki67-Working Group (IKWG) gegründet, die sich mit der Entwicklung von Richtlinien und der Standardisierung des Umgangs mit Ki67 befasst.

1.1.5.1 Ki67 in der Klinik

Ki67 wird vor allem in der wissenschaftlichen Forschung eingesetzt, wobei die zentrale Frage häufig darin besteht, ob Ki67 einen prognostischen Wert für den klinischen Verlauf oder eine bessere Differenzierung von Tumorerkrankungen aufweist. Die Forschungsergebnisse variieren dabei erheblich.

In einigen Fällen konnten bereits signifikante Ergebnisse erzielt werden, die Ki67 als prognostischen Marker für eine schwerere Erkrankung und einen ungünstigeren klinischen Verlauf beschreiben, wie bei neuroendokrinen Neoplasien der Lunge (Naheed et al., 2022) und Prostatakarzinomen (Richardsen et al., 2017). Ebenso wurde festgestellt, dass Ki67 in Bezug auf maligne und benigne Pheochromozytome und Paragangliome als ergänzende Diagnostik bei der Identifizierung von Fällen dienen kann, die ein geringeres Risiko von Rezidivbildung und Metastasierung aufweisen (Elder et al., 2003). Allerdings gibt es auch wiederum Studien, beispielsweise zu kleinzelligen Lungenkarzinomen und extrapulmonaren neuroendokrinen Tumoren (de M Rêgo et al., 2017), Magenkarzinomen (Böger et al., 2015) oder Glioblastomen (Dahlrot

et al., 2021), die keinen prognostischen oder prädiktiven Wert für Ki67 feststellen können.

Besonders im Zusammenhang mit dem Mammakarzinom hat Ki67 Einzug in die klinische Praxis gefunden. Als prognostischer Marker für den Verlauf dieser Tumorentität wurde er in den Leitlinien festgehalten. Allerdings wurde seine Aussagekraft in den vergangenen Jahren immer wieder reevaluiert. Während anfangs ein hoher Ki67-Proliferationsindex mit einem schlechteren Überleben von allen Patienten korreliert wurde (de Azambuja et al., 2007), geht man mittlerweile davon aus, dass eine prognostische Aussage nur für luminale Tumoren mit günstigen anatomischen Voraussetzungen (T1-2, N0-1) getroffen werden kann (Nielsen et al., 2021, Liang et al., 2020). Zudem konnte festgestellt werden, dass mit den aktuellen Analysemethoden nur für Proliferationswerte von <10% oder >25% bzw. ≤5% oder ≥30% eine verlässliche Aussage über einen niedrigen bzw. hohen Proliferationsindex getroffen werden kann (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021, Nielsen et al., 2021).

1.1.5.2 Technische Aspekte und Grenzen der Ki67-Auswertung

Als Standardmethode zur Auswertung von Ki67 empfiehlt die IKWG eine globale Auswertung (Leung et al., 2019). Dabei werden vier Bereiche innerhalb des Tumors definiert, die jeweils eine hohe, mittlere, niedrige oder negative Proliferationsrate repräsentieren. Die Anzahl der gezählten Zellen in jedem Bereich wird mit dem Anteil der Fläche dieser Region am Gesamttumor verrechnet (Leung et al., 2019). Lange Zeit war es üblich, sogenannte „Hot Spots“ auszuwerten, was bedeutet, dass etwa 500 Zellen an der Stelle höchster Proliferation analysiert und der Anteil der gefärbten Zellen berechnet wird (Leung et al., 2019). Eine Studie von Shui et al. verglich die beiden Methoden und stellte fest, dass die globale Einschätzung der Proliferationsrate etwas reproduzierbarer ist als die Hot Spot Methode (Shui et al., 2015). Robertson et al. kamen zu ähnlichen Schlussfolgerungen und fanden, dass sich bei Einsatz der globalen Methode eine bessere Korrelation mit dem Gesamtüberleben zeigt (Robertson et al., 2020).

Die größte Problematik liegt jedoch in der hohen Inter- und Intraobserver-Variabilität im mittleren Proliferationsbereich. Mehrere Studien haben gezeigt, dass in diesem Bereich eine geringe Korrelation zwischen verschiedenen bewertenden Personen festgestellt werden kann (Mikami et al., 2013, Shui et al., 2015) Für die Werte <5% und >30% wurde hingegen eine hohe Korrelation festgestellt (Shui et al., 2015).

Ein weiterer Aspekt ist die Variabilität der Bewertung innerhalb und zwischen verschiedenen Laboren. Vor der Einführung der Protokolle zur Standardisierung durch die IKWG war die Korrelation zwischen den an der Studie teilnehmenden Institutionen sehr gering (Nielsen et al., 2021). Vor allem durch die Einführung eines Protokolls zur globalen Auswertung von Stanzbiopsien konnte im Laufe der Zeit eine deutliche Verbesserung erzielt werden (Nielsen et al., 2021).

Wie bereits erwähnt, wird Ki67 als unabhängiger prognostischer Faktor nur noch für anatomisch günstige luminale Tumoren in Betracht gezogen. Eine Studie von Liang et al. konnte keinen signifikanten Zusammenhang zwischen Ki67 und dem rezidivfreien Überleben des Patienten bei HER2+- und TN-Tumoren feststellen (Liang et al., 2020).

Die aufgeführten Sachverhalte führen dazu, dass die IKWG den prognostischen Wert für Ki67 als begrenzt betrachtet (Nielsen et al., 2021).

1.1.6 Der Mitoseindex

1.1.6.1 Grenzen der Mitosezählung anhand von Hämatotoxylin-Eosin-Färbung (HE)

Der Goldstandard der Mitosebewertung zur Einschätzung des Gratings von Mammakarzinomen besteht nach wie vor im Zählen von Mitosefiguren in 10 HPF anhand der HE-Färbung. Frühzeitig wurde erkannt, dass der MAI einen Einfluss auf das Grading von Mammakarzinomen hat (Bloom and Richardson, 1957, Elston and Ellis, 1991). Insbesondere wird dem MAI ein starker prognostischer Wert für das krankheitsfreie Überleben zugeschrieben (Klintman et al., 2013). Durch strikte Protokolle konnte eine hohe Reproduzierbarkeit der Mitosebewertung erreicht werden (van Diest et al., 1992).

Trotzdem ist die Mitosezählung anhand von HE-Schnitten nicht besonders beliebt unter Patholog:innen. Häufig genannte Probleme sind Schwierigkeiten bei der Unterscheidung von Mitosen zu Apoptosen, Artefakten oder anderen Zelltypen (Skaland et al., 2007, Tapia et al., 2006, Puripat and Loharamtaweethong, 2019). Um diese Herausforderung zu optimieren, wird aktuell an IHC-Färbungen mit mitosespezifischen Antikörpern wie Phosphohiston 3 (PhH3) geforscht, da diese möglicherweise einen positiven Einfluss auf die Auswertung von Mitosefiguren haben könnten.

1.1.7 Der Mitosemarker PhH3

1.1.7.1 Technische Aspekte und Grenzen der PhH3-Auswertung

PhH3 ist ein nukleäres Protein, das zu den vier Kernhistonen gehört. Die Phosphorylierung erfolgt an den Resten Serin 10 (Ser10) oder Ser28 und korreliert mit dem Eintritt der Zelle in die Mitose (Puripat and Loharamtaweethong, 2019). Während der Prophase, Metaphase, Anaphase und Telophase, wenn die Chromosomen stark kondensiert sind, kann es in diesem phosphorylierten Zustand vorgefunden werden. In der Interphase erscheint es hingegen unphosphoryliert (GURLEY et al., 1978, Puripat and Loharamtaweethong, 2019).

PhH3 wurde im Zusammenhang mit verschiedenen Tumorentitäten wie neuroendokrinen Tumoren des Pankreas (Villani et al., 2016), Mantelzelllymphomen (Medani et al., 2021), Meningiomen (Puripat and Loharamtaweethong, 2019, Ribalta et al., 2004) und Melanomen (Nasr and El-Zammar, 2008, Ladstein et al., 2012, Tetzlaff et al., 2013) untersucht. In allen Fällen wurde festgestellt, dass die IHC mit PhH3 das Zählen von Mitosen deutlich erleichtert (Villani et al., 2016, Medani et al., 2021, Puripat and Loharamtaweethong, 2019, Ribalta et al., 2004, Nasr and El-Zammar, 2008, Ladstein et al., 2012, Tetzlaff et al., 2013). Eine Metaanalyse von Hao et al. hat gezeigt, dass eine hohe Expression von PhH3 mit einem schlechteren Gesamtüberleben der Patienten assoziiert werden kann und auch mit einem schlechteren rezidivfreien und krankheitsfreien Überleben zusammenhängt (Hao Q, 2018).

Die Anwendung von PhH3 wurde auch beim Mammakarzinom untersucht. Dabei konnte festgestellt werden, dass der PhH3-Antikörper ein sensitiver Marker für Mitosen ist (Mirzaian et al., 2020, Kim et al., 2017, Ibrahim et al., 2023, Woo et al., 2015) und außerdem eine hohe Übereinstimmung zwischen verschiedenen Pathologen bei der Auswertung mittels IHC erzielt wird (Skaland et al., 2007, Kim et al., 2017). Zusätzlich ermöglicht die IHC-basierte Auswertung eine deutlich schnellere Analyse im Vergleich zur herkömmlichen HE-Methode (Kim et al., 2017, Bosch et al., 2017) (die Studie von Bosch et al. befasst sich mit duktalem Carcinoma in situ (DCIS)). PhH3 erwies sich auch als hilfreich bei der Einschätzung der am stärksten proliferierenden Tumorfläche (Mirzaian et al., 2020, Kim et al., 2017).

Darüber hinaus wurde PhH3 ebenfalls als eigenständiger prädiktiver Faktor für das Gesamtüberleben und das krankheitsfreie Überleben identifiziert (Ibrahim et al., 2023, Kim et al., 2017). In einer Studie von Skaland et al. wurde außerdem festgestellt, dass eine Mitoserate von unter 13 Mitosen pro 1,59 mm² mit einem besseren rezidivfreien Überleben in Bezug auf Fernmetastasen innerhalb eines klinischen Follow-Ups von 20 Jahren korreliert (Skaland et al., 2007).

In Bezug auf Ki67 ergaben zwei Studien eine lineare Korrelation (Kim et al., 2017, Bosch et al., 2017). Gleichzeitig wird jedoch angenommen, dass PhH3 Ki67 und auch den MAI als Marker für Tumorproliferation übertrifft (van Steenhoven et al., 2020, Kim et al., 2017).

Es fällt auf, dass mit der PhH3-IHC-Methode mehr Mitosen erfasst werden als im HE-Schnitt (Ibrahim et al., 2023, Skaland et al., 2007, Mirzaian et al., 2020, Dessauvage et al., 2015). Dies führt häufig zu einer Neubewertung des Tumorgradings (Re-Grading), wenn man sich nach den Regeln der herkömmlichen 10 HPF-Methode richtet (Ibrahim et al., 2023, Woo et al., 2015, Mirzaian et al., 2020, van Steenhoven et al., 2020). Allerdings wurden in verschiedenen Studien unterschiedliche Cut-Off-Werte festgelegt, wie z.B. 13 Mitosen pro 1,59 mm² (Skaland et al., 2007), 13 Mitosen pro 2 mm² (van Steenhoven et al., 2020), oder 20 pro 3 mm² (Ibrahim et al., 2023).

1.1.8 Bisherige multiparametrische Ansätze zur Verbesserung des Gradings vom invasiven Mammakarzinom NST

Die molekulare Klassifikation des invasiven Mammakarzinoms NST, in Verbindung mit dem histologischen Grading, hat sich als bisher effektivster Ansatz zur präzisen Kategorisierung dieser Tumorentität bewiesen. Insbesondere die Gruppe der G2-Tumoren kann hierdurch exakter differenziert werden. Zuvor wurden G2-Tumoren als äußerst heterogene Entitäten betrachtet, was die eindeutige Festlegung von Therapieentscheidungen, wie beispielsweise die Gabe einer zytotoxischen Chemotherapie, erschwerte (Rakha et al., 2010).

Ungeachtet dieser Fortschritte in der Klassifizierung besteht weiterhin das Bestreben, noch differenziertere und präzisere Klassifizierungsmethoden zu entwickeln. Die bislang als Goldstandard geltende morphologische Kategorisierung, die auf histologischen Merkmalen basiert, wird in diesem Kontext von einigen Experten kritisch betrachtet. Diese Methodik scheint die biologische Variabilität der Tumorentitäten unzureichend widerzuspiegeln (Sun et al., 2014). In Anbetracht dessen wurde dokumentiert, dass morphologisch ähnliche Tumoren unterschiedlich auf therapeutische Interventionen ansprechen und signifikante Unterschiede im Rezidivrisiko sowie Gesamtüberleben aufweisen können (Sun et al., 2014, Rakha et al., 2010, Yu et al., 2021).

Die Einführung multiparametrischer Ansätze eröffnete eine vielversprechende Option zur weiteren Verfeinerung des Gradings. Es wurden diverse Klassifikationsmodelle entwickelt, welche individuelle morphologische und molekulare Tumoreigenschaften kombinieren. Als zu berücksichtigende Parameter wurden Faktoren wie das Patientenalter, ethnische Herkunft, Hormonrezeptorstatus (ER, PR, HER2), Tumolvolumen, Lymphknotenstatus, Grading, histologischer Subtyp, Ki67-Index (Yu et al., 2021, Li et al., 2021), makro- und mikrokollagene Signaturen im Tumorgewebe (Xi et al., 2021), Tumordinfiltration und Karzinomstammzellpräsenz (Xiang et al., 2022) in Betracht gezogen. Die vorliegenden Forschungsarbeiten kommen zu dem Schluss, dass die entwickelten multiparametrischen Modelle eine robuste Korrelation zu rezidivfreiem Überleben sowie Gesamtüberleben aufweisen (Li et al., 2021, Xiang et al., 2022, Yu et al., 2021). Sie erweisen sich auch als vielversprechend,

um Prädiktionen im Hinblick auf das Grading und den molekularen Subtyp der Tumoren zu ermöglichen (Xi et al., 2021, Xu et al., 2022, Ohashi et al., 2023).

1.2 Digitale Analyse (DA) und künstliche Intelligenz (KI)

In der modernen Pathologie gewinnen digitale Hilfsmittel zunehmend an Bedeutung, insbesondere vor dem Hintergrund immer spezifischerer, personalisierter Therapiemethoden. Die Einführung moderner Biomarker und molekularer Analysen hat dazu geführt, dass die alleinige visuelle Auswertung von pathologischen Befunden zunehmend herausfordernd wird (Baxi et al., 2022).

Ein großer Vorteil pathologischer Expert:innen ist ihre Fähigkeit, Zellen unter verschiedenen Färbebedingungen subjektiv zu unterscheiden. Diese Subjektivität birgt jedoch auch Schwächen, da die Übereinstimmung in Inter- und Intraobserver-Vergleichen zwar hoch, aber niemals perfekt ist (s.o.). In Zukunft könnten speziell trainierte KIs und Deep-Learning-Algorithmen zu reproduzierbareren und präziseren Diagnosen führen. In der Literatur wurde bereits gezeigt, dass dies bei Klassifikation, Grading und Staging von Mammakarzinomen möglich ist (Wang et al., 2022, Zhu et al., 2022). Ein solcher Fortschritt würde sich auch auf die klinische Praxis auswirken, beispielsweise durch eine optimierte prädiktive und prognostische Entscheidungsfindung bei Krebstherapien oder personalisierte Therapiegestaltung (Baxi et al., 2022, Zhu et al., 2022, Jiang et al., 2020). Zudem sehen einige Forschende die Chance, ein tieferes Verständnis für die Komplexität und die molekularen Mechanismen der Tumormikroumgebung zu entwickeln (Baxi et al., 2022).

Auf ökonomischer Ebene bietet die Einführung von KI die Möglichkeit, große Datenmengen effizient zu verarbeiten und gleichzeitig Infrastruktur besser auszunutzen, was zu Kosteneinsparungen führen kann (Zhu et al., 2022, Jiang et al., 2020, Baxi et al., 2022). Dies bringt u.a. Vorteile für die Entwicklung neuer Medikamente mit sich (Baxi et al., 2022).

Obwohl die Einführung von KI vielversprechende Ergebnisse in der Diagnostik zeigt, ergeben sich gleichzeitig Herausforderungen, beispielsweise in der

Interpretation der Befunde. Nicht immer ist für den Menschen nachvollziehbar, wie KI Entscheidungen trifft (Baxi et al., 2022, Jiang et al., 2020, Zhu et al., 2022).

Darüber hinaus bieten große Datensätze zwar einen enormen Vorteil, aber sie müssen auch verarbeitet werden. Themen wie Speicherkapazitäten und digitale Infrastruktur stellen weiterhin eine Herausforderung für Forschungsinstitute und Kliniken dar, da sie sowohl Kosten als auch Raum beanspruchen (Zhu et al., 2022, Jiang et al., 2020, Baxi et al., 2022). Es kann auch zu Problemen kommen, wenn neue Daten in benutzerdefinierte KIs eingespeist werden (Bankhead, 2022). Daher ist eine enge Zusammenarbeit zwischen Softwareentwickler:innen und Patholog:innen von großer Bedeutung (Bankhead, 2022).

Da KI bisher nur im Studienkontext untersucht wurde, sind weitere Forschung und Validierung erforderlich, bevor diese Technologie endgültig in den klinischen Alltag integriert werden kann (Zhu et al., 2022, Jiang et al., 2020).

Bereits heute findet eine Teilanwendung von KI in der Pathologie statt - die digitale Bildanalyse. Mit der Einführung von Scannern, die Whole Slide Images (WSI) erstellen, werden digitale histologische oder pathologische Bilder in Lehre und Forschung immer häufiger verwendet (Bankhead, 2022, Barisoni et al., 2020, Pell et al., 2019). Die Verbindung von analoger und digitaler Pathologie hat einen großen Markt an kommerziellen und frei zugänglichen Analysesoftwaren wie ImageJ oder QuPath entstehen lassen (Pell et al., 2019, Bankhead, 2022).

Diese Technologie findet insbesondere Anwendung bei der IHC-Analyse von Hormonrezeptoren oder der Bestimmung des prozentualen Anteils gefärbter Zellen im Verhältnis zum Rest des Tumors, wie beispielsweise bei der Auswertung des Ki67-Proliferationsindex in Mammakarzinomen (Acs et al., 2019, Pell et al., 2019). Die digitale Analyse erweist sich als vorteilhaft, da sie die Genauigkeit verbessert und die Interobserver-Variabilität reduziert (Pell et al., 2019). Zudem wurde beschrieben, dass die qualitative Bewertung von Expressionsmustern von Biomarkern oder Gewebetypen in eine quantitative Analyse umgewandelt werden kann. (Aeffner et al., 2019).

Trotz der aussichtsreichen Studienresultate befinden sich diese Hilfsmittel noch in einer Entwicklungsphase und sollten bei der Diagnosefindung und -entscheidung lediglich als Ergänzung zur Arbeit erfahrener Patholog:innen genutzt werden (Bankhead, 2022).

1.2.1 QuPath

QuPath ist eine der frei zugänglichen Bildanalysesoftware, die im Jahr 2017 von Forschenden des Northern Ireland Molecular Pathology Laboratory, Centre for Cancer Research and Cell Biology der Queen's University in Belfast entwickelt wurde. Es handelt sich dabei um eine plattformübergreifende Java-Anwendung, deren Hauptfunktionen das Einlesen und Anzeigen von WSI sowie die Zelldetektion sind (Bankhead et al., 2017). Die Software wurde ursprünglich mit IHC-gefärbten Schnitten trainiert, wobei die verwendeten Biomarker CD3, CD8, p53 und PD-L1 waren (Bankhead et al., 2017).

Aufgrund des kontinuierlichen Austauschs der Entwickler:innen mit den Nutzer:innen über Onlineforen befindet sich die Plattform in ständiger Weiterentwicklung. Abbildung 2 zeigt die Fachbereiche, in denen QuPath im klinischen Alltag am häufigsten genutzt wird:



Abbildung 2: Übersicht der fünf Fachbereiche, die QuPath am häufigsten nutzen. Die Relevanz wurde durch Humphries et al. anhand der Anzahl wissenschaftlicher Veröffentlichungen gemessen, die das ursprüngliche QuPath-Paper zitieren (Humphries et al., 2021)

QuPath entfaltet sein größtes Potential bei der Erkennung von IHC-gefärbten Zellen und der Differenzierung von umgebendem Gewebe wie Stroma oder Artefakten. Mithilfe des „Positive Cell Detection“-Tools kann es den Anteil der gefärbten Zellen im Verhältnis zur Gesamtzahl der Tumorzellen berechnen.

Allerdings werden auch Schwierigkeiten bei der Analyse von IHC-Markern mit niedrigem klinischen Cut-Off angegeben (Bankhead et al., 2017).

1.2.2 Digitale Analyse von Ki67 mit Hilfe von QuPath

Unter den verschiedenen Anwendungen zur digitalen Auswertung von Biomarkern wird insbesondere die digitale Bewertung des Proliferationsindex beim Mammakarzinom intensiv erforscht. Im Jahr 2018 entwickelte die IKWG ein Protokoll, dass die Analyse von Ki67-gefärbten Whole Slide Images (WSI) mithilfe von QuPath für den klinischen Alltag nutzbar und standardisiert macht (Acs et al., 2019, Acs et al., 2022). In mehreren Studien konnte eine gute Reproduzierbarkeit der Methode festgestellt werden und sie scheint den konventionellen Methoden ebenbürtig zu sein (Acs et al., 2019, Nielsen et al., 2021, Boyaci et al., 2021).

Darüber hinaus wurde der digitale globale Score als unabhängiger Prognosefaktor für das Gesamt- und rezidivfreie Überleben bei luminalen Tumoren identifiziert (Robertson et al., 2020). Robertson et al. fanden außerdem heraus, dass die digitale globale Auswertung im Durchschnitt einen geringeren Proliferationsindex ergibt (15,9% im Vergleich zu 20,0% bei der manuellen Auswertung) und dass die digitale Methode der manuellen Auswertung überlegen zu sein scheint (Robertson et al., 2020). Unterschiede wurden auch beim Vergleich verschiedener pathologischer Präparate festgestellt, wobei der Proliferationsindex von Stanzbiopsien im Durchschnitt höher ausfällt als der von Resektaten. Dies könnte mit der Behandlung des Gewebes oder der Fixierung zusammenhängen (Acs et al., 2022). Diese Erkenntnisse bestätigen den allgemeinen Konsens der IKWG, wonach vor allem die Stanzbiopsien für die Bewertung von Ki67 im klinischen Alltag herangezogen werden sollten (Acs et al., 2022).

Obwohl die digitale Analyse von Ki67 mit QuPath vielversprechende Ergebnisse zeigt, gibt es bisher nur wenige unabhängige Studien zu diesem Thema. Eine Arbeit von Skjevold et al. beschreibt beispielsweise, dass bei der digitalen Analyse von Ki67 mehr Fälle mit hohem Proliferationsindex ausgewertet werden als bei der visuellen Methode. Außerdem geben sie an, dass die digitale Methode unpraktikabel im klinischen Alltag ist (Skjevold et al., 2022). Im Gegensatz dazu bestätigen Boyaci et al. den prognostischen Wert des Protokolls zur digitalen Analyse mit QuPath, wie von der IKWG entwickelt (Boyaci et al., 2021).

Angesichts der begrenzten Anzahl unabhängiger Studien ist es notwendig, den Effekt der digitalen Analyse auf die Auswertung des Ki67-Proliferationsindex weiter zu evaluieren.

1.2.3 Digitale Analyse von PhH3

Die digitale Analyse des Mitoseindex befindet sich noch in einem vergleichsweise wenig etablierten Stadium. Bisher gilt die zeitaufwändige visuelle Zählung von Mitosefiguren am HE-Schnitt als Goldstandard. Dennoch werden zunehmend Bemühungen unternommen, diese Methode zu automatisieren.

Die bisherigen Studien zur digitalen Mitosebewertung umfassen verschiedene Tumorentitäten wie Brustkrebs (Dessauvagie et al., 2015, Ibrahim et al., 2023, Skaland et al., 2007), gastrointestinale Tumoren (Jin et al., 2019), Melanome (Hale et al., 2013, Nielsen et al., 2016), oder Mantelzelllymphome (Medani et al., 2021). Um die Mitosefiguren besser für Bildanalysesoftwares sichtbar zu machen, wird in all diesen Studien der bereits erwähnte Mitosemarker PhH3 eingesetzt (Nielsen et al., 2016, Medani et al., 2021, Ibrahim et al., 2023, Hale et al., 2013, Jin et al., 2019, Dessauvagie et al., 2015, Skaland et al., 2007). Die Umsetzung der einzelnen Studien ist ähnlich, dennoch lassen sich Unterschiede feststellen:

Für die Auswertung werden verschiedene Software-Tools verwendet, darunter das Visiomorph DP-Modul (Visopharm) (Nielsen et al., 2016) ImageJ und/oder QuPath (Ibrahim et al., 2023, Skaland et al., 2007), Aperio ScanScope (Hale et

al., 2013, Dessauvagie et al., 2015), Fiji Software (Medani et al., 2021) und Ventana Virtuoso (Jin et al., 2019).

In allen Studien wurden als auszuwertende Regionen Hot Spots definiert. Diese wurden manuell oder digital dort festgelegt, wo die größte Anzahl an PhH3-gefärbten Zellen festgestellt werden konnte (Hale et al., 2013, Ibrahim et al., 2023, Nielsen et al., 2016, Jin et al., 2019). Laut Skaland et al. befinden sich diese Hot Spots vor allem in der peripheren Zone des Tumors (Skaland et al., 2007). Die Fläche der ausgewerteten Regionen variiert dabei von 0,5-3 mm² (Ibrahim et al., 2023, Nielsen et al., 2016, Hale et al., 2013, Jin et al., 2019, Dessauvagie et al., 2015). Teilweise wurden auch mehrere HPF ausgewertet, ähnlich der konventionellen Methode (Skaland et al., 2007, Medani et al., 2021). Um ein adäquates Ergebnis zu erzielen, scheint es wichtig zu sein, beim Annotieren der Fläche nur Tumorgewebe zu markieren (Dessauvagie et al., 2015).

Interessanterweise konnte in den meisten Studien herausgefunden werden, dass die digitale Auswertung von PhH3 stark mit der visuellen korreliert (Skaland et al., 2007, Dessauvagie et al., 2015, Ibrahim et al., 2023, Hale et al., 2013). Es gibt aber auch Arbeitsgruppen, die beschrieben, dass die digitale Methode in IHC-Schnitten mehr Mitosen detektiert (Jin et al., 2019). Nielsen et al. stellten in ihrer Studie diesbezüglich fest, dass der Mitoseindex bei der Verwendung der automatisierten Zählung im Durchschnitt höher ausfällt, die Gesamtzahl der detektierten Zellen jedoch ähnlich der der manuellen Zählung ist (Nielsen et al., 2016). In Bezug auf die digitale Auswertung von HE-Schnitten fällt die dort detektierte Mitosezahl, ähnlich wie bei der visuellen Auswertung, deutlich geringer aus als bei der Analyse von PhH3 (Ibrahim et al., 2023).

Einschränkungen der automatischen Methode wurden vor allem bei der Analyse kleiner Tumorflächen, sehr dunkler Färbung oder einer großen Anzahl von Lymphozyten im Probengewebe festgestellt (Nielsen et al., 2016).

In den bisherigen Studien lässt sich keine einheitliche Methode finden, die einen globalen Analyseansatz verfolgt. Dennoch könnte die digitale Evaluation von

Mitosen, insbesondere mit Hilfe einer PhH3 Färbung, eine Erleichterung für die bisher eher mühsame Arbeit der Mitosezählung darstellen.

1.3 Fragestellung und Ziele

Die vorliegenden Erkenntnisse verdeutlichen, dass angesichts der beträchtlichen Anzahl von Patient:innen (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft and AWMF), 2021) und den damit verbundenen ökonomischen Belastungen des Gesundheitssystems (Schöffski and Emmert, 2009) ein großes Interesse darin besteht, die Therapieansätze und Diagnoseverfahren für das Mammakarzinom zu optimieren. Innerhalb dieses Kontextes spielt die visuelle Bewertung des Proliferationsmarkers Ki67 und des Mitoseindex eine entscheidende Rolle. Forscher:innen stoßen wiederholt auf Beschränkungen bei diesen Verfahren, die potenziell erheblichen Einfluss auf die Therapieentscheidungen beim Mammakarzinom haben können.

Insbesondere Interobserver-Variabilitäten führen dazu, dass für die Einschätzung der Proliferationsindizes von Mammakarzinomen nur vage Aussagen getroffen werden können, insbesondere wenn es sich um Tumoren mit moderater Proliferation (im Bereich von 5-30%) handelt (Nielsen et al., 2021). Aufgrund der Assoziation einer hohen Expression dieses Markers mit einem geringeren rezidivfreien Überleben, vor allem im Kontext mit luminalen HER2 negativen Tumoren (Liang et al., 2020), besteht ein dringender Bedarf an präzisen und reproduzierbaren Analysemethoden. In diesem Zusammenhang scheinen sich standardisierte Protokolle und digitale Hilfsmittel als nützlich zu erweisen (Robertson et al., 2020).

Des Weiteren bieten sich Optimierungspotenziale im Bereich des morphologischen Gradings, speziell bei der Zählung von Mitosefiguren. Hier könnte eine IHC-Färbung mit dem Marker PhH3 den Zeitaufwand im Vergleich zur konventionellen Auswertung der HE Färbung erheblich reduzieren (Ibrahim et al., 2023). Digitale Analysen könnten ebenfalls dazu beitragen Interobserver-Variabilitäten zu minimieren und den Prozess der Mitosezählung weiter zu standardisieren.

Studien haben gezeigt, dass multiparametrische Klassifikatoren eine verbesserte prädiktive Aussagekraft hinsichtlich des rezidivfreien Überlebens und des Gesamtüberlebens aufweisen können, was eine präzisere Einteilung der Tumoren ermöglicht (Li et al., 2021). Angesichts dieser Erkenntnis stellt sich die Frage, ob die Integration weiterer Parameter einen ähnlichen Einfluss auf den klinischen Verlauf haben kann.

Zwischen PhH3 und dem klinischen Verlauf des Mammakarzinoms wurde eine starke Korrelation festgestellt (Hao Q, 2018). Es konnte sogar nachgewiesen werden, dass PhH3 als unabhängiger Prädiktionsfaktor für das Gesamtüberleben fungiert (Kim et al., 2017). Ebenso wurde ein linearer Zusammenhang mit dem Ki67-Proliferationsindex aufgezeigt (Bosch et al., 2017). Vor dem Hintergrund dieser Resultate stellt sich die Frage, ob die Berücksichtigung beider Biomarker für die molekulare Klassifikation einen Einfluss auf den klinischen Verlauf haben könnte.

Das Hauptziel dieser Arbeit besteht darin, den potenziellen Einfluss eines multiparametrischen Ansatzes, welcher visuelle und digitale Analysen von Ki67 und PhH3 kombiniert, auf das konventionelle Grading und den klinischen Verlauf vom invasiven Mammakarzinom NST zu untersuchen. Dabei sollen die verschiedenen Analysemethoden miteinander verglichen werden, um festzustellen, inwiefern sie sich bei der Bestimmung des Proliferations- und Mitoseindex, auch in Bezug auf unterschiedliche Probenarten (Resektat und Stanzbiopsien), voneinander unterscheiden. Ein weiteres Ziel ist die Untersuchung, ob ein digitaler Ansatz Möglichkeiten zur Optimierung der Einzelparameterbestimmung bietet. Darüber hinaus ist beabsichtigt, langfristig eine Kohorte zu etablieren, die als Goldstandard für künftige Forschungsarbeiten, welche auf den Einsatz von KI-basierten Methodologien abzielen, fungieren soll.

2 Material und Methoden

2.1 Die Kohorten

In dieser Studie wurden zwei Kohorten ausgewertet. Die erste Kohorte, die im Rahmen eines vorherigen Projektes zusammengestellt wurde, dient als Entdeckungskohorte für die vorliegende Arbeit. Eine zweite Kohorte wurde auf der Grundlage engerer Auswahlkriterien unter Berücksichtigung der klinischen Daten erstellt und fungiert als Validierungskohorte für die Erkenntnisse der ersten Kohorte. Die Verwendung der Proben erhielt am 19.08.2021 die Zustimmung der Ethik-Kommission der Universität Tübingen (Projektnummer: 547/2021BO2).

Die Initialauswahl der Fälle erfolgte mit Hilfe von tabellarischen Excel-Dateien aus der Datenbank der strukturierten Mammakarzinombefunde des Pathologischen Institutes, welche die Mammakarzinomfälle des Uniklinikum Tübingen jährlich zusammenfassen. Diese Datensätze umfassten eine beträchtliche Anzahl von über 2000 Fällen und mussten zunächst anhand von Kriterien für NST-Tumoren und spezifischer Einschlusskriterien der einzelnen Kohorten (weiter unten beschrieben) gefiltert werden.

Die pathologischen Schnitte, die zu den vorausgewählten Fällen gehörten, wurden aus dem Archiv des Instituts für Pathologie und Neuropathologie gesammelt. Im Anschluss erfolgte eine erneute Evaluierung der HE-Präparate durch erfahrene Pathologinnen des Instituts (Prof. Dr. Annette Staebler, Dr. Ivonne Montes). Diese Überprüfung zielte darauf ab, die Qualität der Präparate zu sichern und sicherzustellen, dass ausreichend Tumormaterial für die geplante digitale Auswertung vorhanden war.

Für alle eingeschlossenen Fälle wurden die entsprechenden, in Formalin fixierten und Paraffin eingebetteten (FFPE-)Gewebeblöcke der Stanzbiopsien bzw. Resektate gesammelt und neue HE-Schnitte angefertigt. Im Einklang mit dem aktuellen Standard, dem Nottingham-Grading-System (Galea et al., 1992), wurden die Fälle erneut graded. Ebenso wurden mehrere IHC-Untersuchungen durchgeführt und ausgewertet (siehe unten).

2.1.1 Entdeckungskohorte

Die Entdeckungskohorte umfasst insgesamt 72 Fälle von FFPE Stanzbiopsien der Mamma. Diese Fälle stammen aus den Jahren 2017 und 2018 und wurden nach zwei Hauptkriterien ausgewählt: das weibliche Geschlecht sowie das Vorliegen eines invasiven Mammakarzinoms NST.

Die Grading-Verteilung innerhalb dieser Kohorte zeigt, dass insgesamt 61% der Fälle, laut ursprünglichem Befund, ein Grading von 3 aufweisen und somit die größte Gruppe bilden. Es ist zu beachten, dass diese Kohorte in erster Linie als Lernkohorte für die digitale Analyse diente. Daher wurde sie so zusammengestellt, dass jede molekulare Subgruppe des invasiven Mammakarzinoms NST in etwa gleiche Fallzahlen aufweist. Die genaue Aufschlüsselung bezüglich der Grade und molekularen Subtypen kann in Tabelle 2 eingesehen werden:

| | Fallzahl (n) | Anzahl der Fälle (%) |
|----------------------------|--------------|----------------------|
| Gesamt | 72 | 100 |
| Histologischer Grad | | |
| Grad 1 | 12 | 17 |
| Grad 2 | 16 | 22 |
| Grad 3 | 44 | 61 |
| Molekularer Subtyp | | |
| HR+ HER2- (Lum. A) | 15 | 21 |
| HR+ HER2- (Lum. B) | 14 | 20 |
| HR+ HER2+ (Lum. B) | 18 | 25 |
| HER2+ HR- | 11 | 15 |
| TN | 14 | 19 |

Tabelle 2: Zusammensetzung des Patientenkollektivs der ersten Kohorte. Die größte Gruppe besteht aus Fällen mit Grad 3. Die einzelnen Subtypen verteilen sich etwa gleichmäßig, wobei die HER2+ HR+ Gruppe mit 25% den größten Anteil ausmacht.

2.1.2 Validierungskohorte

Die Validierungskohorte wurde in Hinblick auf die Korrelation mit dem klinischen Verlauf nach strengeren Einschlusskriterien erstellt. Es wurden lediglich Fälle einbezogen, bei denen die FFPE-Stanzbiopsien und Resektate a priori verfügbar waren. Ferner musste die Operation (Mastektomie, BET) in domo durchgeführt worden und keine neoadjuvanten Behandlungen erfolgt sein. Darüber hinaus wurden nur primäre, unifokale Tumoren ohne Lymphknotenmetastasen (N0) oder

Fernmetastasen (M0) berücksichtigt. Männliche Patienten wurden ebenfalls exkludiert (n=2).

Um eine sinnvolle Korrelation der erhobenen Daten mit dem klinischen Verlauf zu ermöglichen, wurden konsekutive Fälle zu den jeweiligen Subtypen aus dem Jahr 2012 in die Kohorte aufgenommen, um ein 10-Jahres-Follow-Up abdecken zu können. Die Serie wurde ergänzt durch Fälle aus den Jahren 2010 und 2011 für die Kategorie HER2+ und TN. Weil diese Patientinnen häufig neoadjuvante Chemotherapien erhalten, waren Resektate ohne Vorbehandlung im Jahr 2012 unterrepräsentiert. Das Hauptziel dieser Validierungskohorte ist es, die Ergebnisse der Entdeckungskohorte zu bestätigen und insbesondere den Einfluss des multiparametrischen Analyseansatzes auf Fälle zu beleuchten, die sich in Bezug auf ihren Proliferationsindex oder ihrem Mitosescore in einem therapeutisch relevanten Grenzbereich befinden. Dies betrifft insbesondere die luminalen Tumoren mit einem Grad 2. Diese Kohorte umfasst 21 lumbale Tumoren mit Grad 1, 42 mit Grad 2 und 22 mit Grad 3. Zusätzlich wurden 21 Fälle mit HER2-Expression und 25 TN-Fälle in die Studie aufgenommen, um auch den Einfluss auf nicht-luminale Subtypen zu analysieren. Eine detaillierte Übersicht dieser Kohorte ist in Tabelle 3 einzusehen.

Anhand der TNM-Klassifikation wird der Großteil dieser Kohorte als pT1c klassifiziert (53%). Die zweitgrößte Gruppe bilden Fälle mit einer Einstufung als T2 Tumor (30,5%).

Die molekulare Klassifikation erfolgte auf Grundlage der FFPE-Stanzbiopsien und orientierte sich an der Re-Evaluation, da in den Jahren der Diagnosestellung nur wenige Fälle über einen definierten Ki67-Wert verfügten.

Wie aus Tabelle 3 ersichtlich ist, sind reine HER2+ Fälle in geringer Zahl vertreten. Dies ist darauf zurückzuführen, dass nur HER2+ Fälle in der Regel neoadjuvant behandelt werden oder die Gewebeproben nur minimale Anteile invasiven Tumorgewebes aufweisen (häufig handelt es sich um DCIS-Fälle). Um eine angemessene Anzahl von HER2+ Fällen zu gewährleisten (mindestens 20), wurden hier auch Fälle mit einbezogen, bei denen keine Stanzbiopsie verfügbar

war, oder diese aufgrund der Qualität des Gewebes nicht verwendet werden konnte (n=4).

| | Fallzahl (n) | Anzahl der Fälle (%) |
|----------------------------|--------------|----------------------|
| Gesamt | 131 | 100% |
| Histologischer Grad | | |
| Grad 1 | 21 | 16 |
| Grad 2 | 55 | 42 |
| Grad 3 | 55 | 42 |
| Molekularer Subtyp* | | |
| HR+ HER2- (Lum. A) | 42 | 32 |
| HR+ HER2- (Lum. B) | 42 | 32 |
| HR+ HER2+ (Lum B) | 14 | 11 |
| HER2+ HR- | 3 | 2 |
| TN | 25 | 19 |
| Klinische Daten | | |
| Alter | | |
| >55 | 81 | 62 |
| <55 | 50 | 38 |
| Betroffene Seite | | |
| Links | 56 | 43 |
| Rechts | 75 | 57 |
| Tumorgröße (T) | | |
| T1a | 2 | 1,5 |
| T1b | 15 | 11 |
| T1c | 69 | 53 |
| T2 | 40 | 30,5 |
| T3 | 5 | 4 |
| Tod | | |
| Ja | 19 | 14 |
| Nein | 107 | 82 |
| Fehlende Daten | 5 | 4 |
| Rezidiv | | |
| Ja | 11 | 8 |
| Nein | 107 | 82 |
| Fehlende Daten | 13 | 10 |

Tabelle 3: Zusammensetzung des Patientenkollektivs der zweiten Kohorte. Die größten Gruppen bilden Fälle mit Grading 2 und 3. Die luminalen Fälle (A/B) sind unter den Subtypen am häufigsten vertreten. Die Anzahl der Todesfälle und Rezidive in dieser Kohorte ist eher gering. (*der für die klinische Behandlung relevante molekulare Subtyp wurde an den Stanzbiopsien bestimmt)

2.1.3 Follow-Up Daten

Die Follow-Up Daten wurden in Zusammenarbeit mit der Universitätsfrauenklinik Tübingen von Dr. Léa-Louise Vollmer erhoben und bereitgestellt. Hierzu wurde

eine offizielle Anfrage beim Klinischen Krebsregister des CCCs Tübingen gestellt. Im Rahmen dieser Kooperation wurden uns folgende Daten für diese Studie zur Verfügung gestellt:

- Vorangegangene und nachfolgende Diagnosen in Bezug auf die primäre Diagnose des Mammakarzinoms
- Diagnosedatum
- Datum der Diagnose des ersten Lokalrezidivs
- Art des ersten Lokalrezidivs
- Datum der Diagnose der ersten Metastase
- Datum der letzten Follow-Up-Erhebung
- Datum des Eingangs der letzten bekannten Information bzgl. Der Krebserkrankung
- Sterbedatum
- Daten der Diagnosen aller bekannter Metastasen
- Information über das Auftreten von Rezidiven
- Information über den Status des Versterbens der Patientin
- Anzahl der krankheitsfreien Tage
- Anzahl der Tage des Gesamtüberlebens
- Anzahl der rezidivfreien Monate
- Anzahl der Monate des Gesamtüberlebens.

Aus diesen Daten lässt sich herausstellen, dass die meisten Patientinnen bei ihrer Erstdiagnose älter als 55 Jahre alt waren (62%) (siehe Tabelle 3).

Zudem wurde das rezidivfreie Überleben als der Zeitraum zwischen der Diagnosestellung und dem Auftreten von Lokalrezidiven, Lymphknoten- und Fernmetastasen definiert. Dementsprechend stellt das Gesamtüberleben den Zeitraum zwischen Diagnosestellung und Tod der Patientin dar.

Es ist erwähnenswert, dass die Gesamtzahl der Todesfälle (14%) und der Rezidive (8%) innerhalb unserer Kohortenverteilung vergleichsweise gering ist (siehe Tabelle 3).

2.2 Angewendete Färbeverfahren

2.2.1 HE-Färbung

Zur Gewährleistung einer einheitlichen Auswertung wurden für diese Studie von jedem gesammelten Präparat neue HE-Schnitte angefertigt. Die Durchführung der Färbungen erfolgte durch erfahrene medizinisch-technische Assistent:innen (MTA) im Eingangslabor des pathologischen Instituts Tübingen.

Der Prozess begann mit dem präzisen Zuschnitt der Tumorblöcke auf eine Dicke von 2,5 µm mittels eines Eprexia HM 355S Automated Microtome (Thermo Fisher Scientific). Die Gewebeschnitte wurden auf passende Objektträger aufgetragen (Tabelle 11/12). Für die Färbung selbst wurde der Tissue-Tek Färbeautomat der Firma Sakura verwendet. Hierbei kam ein standardisiertes Protokoll zur Anwendung, welches detailliert in Tabelle 4 aufgeführt ist.

| Arbeitsschritt | Zeit in min |
|-----------------------|--------------------|
| Trockenzeit | 12:00 |
| Xylol | 2:30 |
| Xylol | 2:30 |
| Alkohol 100% | 1:00 |
| Alkohol 96% | 0:30 |
| Alkohol 70% | 0:30 |
| Destilliertes Wasser | 0:30 |
| Hämalaun | 2:00 |
| Hämalaun | 4:00 |
| Hämalaun | 4:00 |
| Leitungswasser | 10:00 |
| Eosin | 0:30 |
| Leitungswasser | 0:10 |
| Alkohol 70% | 0:20 |
| Alkohol 96% | 1:00 |
| Alkohol 100% | 1:00 |
| Alkohol 100% | 1:00 |
| Xylol | 4:00 |

Tabelle 4: Protokoll zur automatisierten HE-Färbung mit dem Tissue-Tek Färbeautomaten (Sakura).

2.2.1.1 Verwendete Reagenzien und Lösungen

2.2.1.1.1 Hämaluan nach Meyer

| Reagenz/Lösung | Menge | Hersteller | Bestellnummer |
|-----------------------|-------|--|---------------|
| Hämatoxylin | 2,5g | Merck KGaA, Darmstadt, Deutschland | 15938 |
| Natriumjodat | 0,5g | Merck KGaA, Darmstadt, Deutschland | 6525 |
| Chloralhydrat | 125g | Sigma-Aldrich, St. Louis, Missouri, Vereinigte Staaten | 15307 |
| Aluminiumkaliumsulfat | 125g | Merck KGaA, Darmstadt, Deutschland | 1047 |
| Zitronensäure | 2,5g | Merck KGaA, Darmstadt, Deutschland | 243 |
| Destilliertes Wasser | 3,5 l | | |

Tabelle 5: Verwendete Reagenzien zur Herstellung der Hämaluanvorratslösung nach Mayer.

2.2.1.1.2 Eosin (1% wässrig)

| Reagenz/Lösung | Menge | Hersteller | Bestellnummer |
|----------------------|-------|------------------------------------|---------------|
| Eosin | 25g | Merck KGaA, Darmstadt, Deutschland | 15935 |
| Destilliertes Wasser | 25 l | | |

Tabelle 6: Verwendete Reagenzien zur Herstellung der Eosinvorratslösung (1% wässrig).

2.2.1.1.3 Weitere Reagenzien

| Reagenz/Lösung | Hersteller |
|------------------|--|
| Xylol | SAV Liquid Production GmbH, Flinzbach am Inn, Österreich |
| Isopropylalkohol | SAV Liquid Production GmbH, Flinzbach am Inn, Österreich |

Tabelle 7: Zusätzliche Reagenzien und Lösungen, die für die Herstellung eines HE-Schnittes notwendig sind.

2.2.2 Immunhistochemische Färbung

Die Immunhistochemischen Schnitte wurden ebenfalls von erfahrenen MTAs des immunhistochemischen Labor des pathologischen Instituts Tübingen (Christiane Stoffregen, Claudia Hermann) hergestellt. Hierfür wurden die FFPE-Blöcke mit dem bereits erwähnten Feinschnitt-Mikrotom (siehe Tabelle 11) ebenfalls auf eine Dicke von 2,5 µm geschnitten. Diese Schnitte wurden anschließend in einem heißen Wasserbad geglättet und danach auf positiv geladene Objektträger aufgetragen. Dort trockneten sie über Nacht bei etwa 37 Grad Celsius.

Analog zu den HE-Schnitten erfolgte die Färbung der immunhistochemischen Schnitte mithilfe eines automatischen Färbeautomaten, dem VENTANA Benchmark ULTRA von Roche (siehe Tabelle 11). Sowohl die Färbung von Ki67 als auch von PhH3 wurde mittels monoklonaler Antikörper durchgeführt (siehe Tabelle 8).

| Antigen | Ki-67 | PhH3 |
|------------------------|---------|---------|
| Antikörper | Mib-1 | phosS10 |
| Färbemuster | nukleär | nukleär |
| Klon-Nr. | M7240 | Ab5176 |
| Hersteller | DAKO | Abcam |
| Verdünnung | 1:400 | 1:500 |
| Inkubationszeit in min | 32 | 32 |
| Inkubationstemperatur | 37 °C | 37 °C |

Tabelle 8: Verwendete Antikörper für Ki67 und PhH3.

Die Färbung basiert auf einem indirekten Färbeverfahren mit dem OptiView Kit der Firma Roche (siehe Tabelle 10). Zunächst erfolgt die Entparaffinierung des

Schnittes mittels EZ-Präp Lösung (Roche) (siehe Tabelle 9). Die Bindungsstellen der Zellen werden mithilfe eines „Cell Conditioner“ (Cell-Conditioner Nr. 1, Roche, Tabelle 9) demaskiert. Nachfolgend wird der Schnitt mit einer dreiprozentigen Wasserstoffperoxidase Lösung (OptiView Peroxidase Inhibitor, Roche, siehe Tabelle 10) inkubiert. Daraufhin wird ein Primärantikörper aufgetragen, der spezifisch an das Zielgen (Ki67, PhH3, siehe Tabelle 8) bindet.

Es folgt das Auftragen eines sekundären Antikörpers (OptiView HQ universal Linker, siehe Tabelle 10), welcher an den Primärantikörper bindet und als Signalverstärker fungiert. Durch die Zugabe und Bindung eines tertiären Antikörpers (OptiView HRP Multimer, Roche) wird durch die Reaktion der enthaltenen Meerrettichperoxidase, einer weiteren 0,4-prozentigen Wasserstoffperoxidlösung (OptiView H₂O₂, Roche), einer Chromogen-Lösung (OptiView DAB, Roche) und Kupfersulfatlösung (OptiView Copper, Roche) ein braun gefärbter Niederschlag erzeugt. Dieser ist später als positive Färbung im IHC-Schnitt erkennbar (Tabelle 10, Abbildung 3A).

Der gesamte Prozess läuft vollautomatisch ab, wobei die Inkubationszeit und -temperatur gemäß den Vorgaben der Antikörperhersteller (siehe Tabelle 8) eingehalten werden. Der BenchMark ULTRA (Roche) wäscht ungebundene Antikörper nach jedem Inkubationsschritt ab und trägt gleichzeitig eine flüssige Deckschicht („liquid cover slip“, LCS, Roche, siehe Tabelle 9) auf, um das Verdunsten wässriger Reagenzien auf dem Objektträger zu verhindern (siehe Abbildung 3B).

Um den Kontrast zu verbessern, wird der Schnitt am Ende des Prozesses mit Hämatoxylin und „blue reagent“ (Roche) gegengefärbt (siehe Tabelle 9). Nach einer 20-minütigen Inkubation wird auch dieser Schnitt mittels des Microm CTM6 (Thermo fisher Scientific) mit einem Deckgläschen versehen.

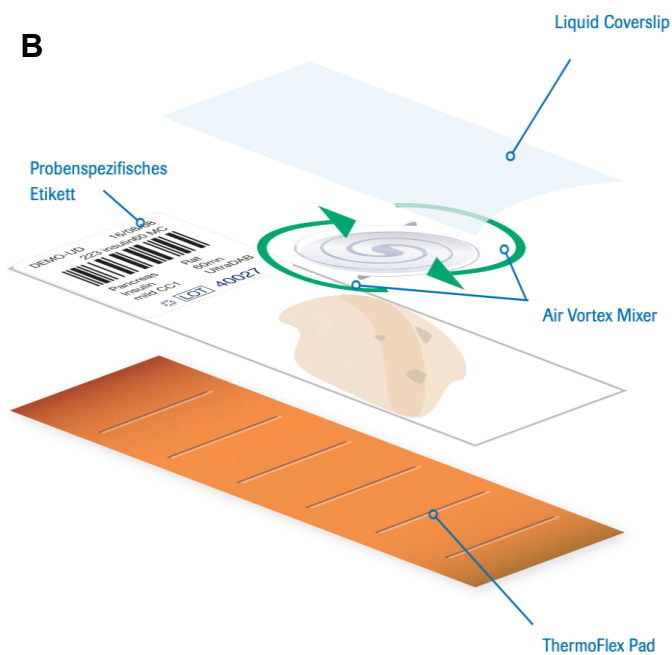
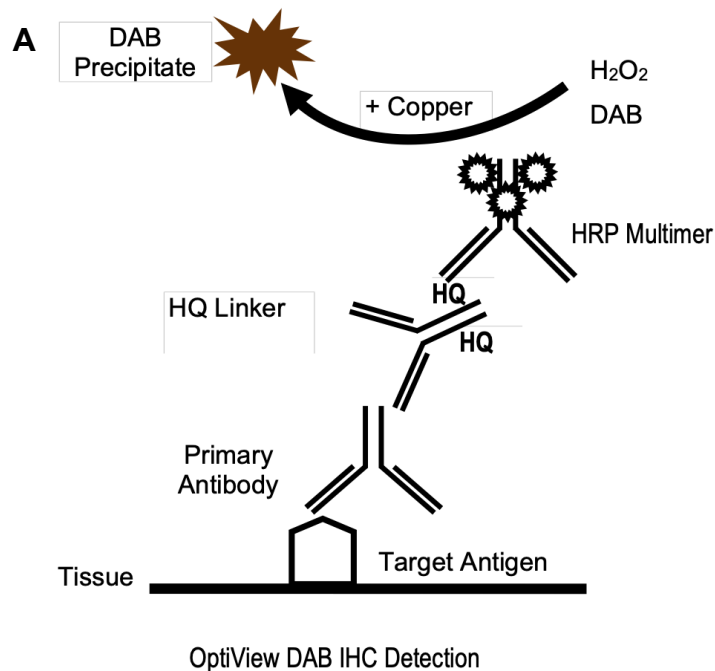


Abbildung 3: Darstellung der Funktionsweisen des OptiView DAB Detection Kit (Roche, 2023a) und des VENTANA BenchMark ULTRA (Roche, 2023b). (A) Die Abbildung stellt schematisch dar, wie die einzelnen Antikörper des Detektionskits an die Zielstruktur binden und mithilfe von weiteren Reagenzien eine immunhistochemische Färbung erzeugen. (B) Diese Abbildung zeigt schematisch die automatisierte Herstellung eines immunhistochemischen Schnittes durch den VENTANA Benchmark ULTRA Färbeautomaten.

2.2.2.1 Verwendete Reagenzien und Lösungen

2.2.2.1.1 Immunhistochemie

| Reagenzien | Hersteller |
|----------------------------------|--|
| OptiView Detection Kit | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| Hämatoxilin | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| Bluing Reagent | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| Reaction Buffer Concentrate | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| Cell Conditioning Solution (CC1) | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| EZ Prep Concentrate | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| LCS (Predilute) | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| Antibody Diluent | Roche Holding AG, Basel/Kaiseraugst, Schweiz |

Tabelle 9: Verwendete Reagenzien für die Immunhistochemischen Färbungen

2.2.2.1.1.1: OptiView Detection Kit (Roche)

| Name | Inhalt |
|-------------------------------|--|
| OptiView Peroxidase Inhibitor | 3,0 % Wasserstoffperoxidlösung. |
| OptiView HQ Universal Linker | HQ-markierte Antikörper (Ziege-Anti-Maus-IgG, Ziege-Anti-Maus-IgM und Ziege-Anti-Kaninchen) (<50 µg/ml) in einem Puffer, der den Konservierungsstoff ProClin 300 enthält. HQ ist ein kovalent an die Ziegenantikörper gebundenes firmeneigenes Hapten. |
| OptiView HRP Multimer | monoklonaler Anti-HQ-markierter HRP Maus-Tertiärantikörper (<40 µg/ml) in einem Puffer, der den Konservierungsstoff ProClin 300 enthält |
| OptiView H2O2 | 0,04 % Wasserstoffperoxid in einer Phosphatpufferlösung. |

| | |
|-----------------|--|
| OptiView DAB | 0,2 % 3,3'-Diaminobenzidin-Tetrahydrochlorid (DAB) in einer firmeneigenen Stabilisierungslösung mit einem firmeneigenen Konservierungsmittel |
| OptiView Copper | Kupfersulfat (5 g/l) in einem Azetat-Puffer mit einem firmeneigenen Konservierungsmittel. |

Tabelle 10: Zusammensetzung des OptiView Deteciton Kit laut Hersteller Roche (Roche, 2023a)

2.2.3 Verwendete Geräte und Verbrauchsmaterialien

| Verwendete Geräte | Hersteller |
|-------------------------------------|--|
| VENTANA BenchMark ULTRA | Roche Holding AG, Basel/Kaiseraugst, Schweiz |
| Tissue-Tek Prima Plus | Sakura Finetek Germany GmbH |
| Microm CTM6 | Thermo Fisher Scientific GmbH, Waltham, Massachusetts, USA |
| Epredia HM 355S Automated Microtome | Thermo Fisher Scientific GmbH, Waltham, Massachusetts, USA |

Tabelle 11: Verwendete Geräte für HE-und immunhistochemische Färbungen.

| Objektträger und Coverslips | Hersteller |
|-------------------------------|--|
| Epredia Coverslips | Thermo Fisher Scientific GmbH, Waltham, Massachusetts, USA |
| Microscope KP FROST SLIDES | Klinipath, Duiven, Niederlande |
| TOMO IHC Adhesive Glass Slide | Matsunami Glass, Osaka, Japan |

Tabelle 12: Verwendete Objektträger und Coverslips

2.3 Visuelle Auswertung

2.3.1 Re-Grading anhand der HE-Schnitte

Das Re-Grading der einzelnen Fälle erfolgte gemäß dem aktuell gültigen Standard der Nottingham-Klassifikation, basierend auf den von Elston und Ellis definierten Parametern (Elston and Ellis, 1991, Galea et al., 1992). Hierbei wurde das Scoring der drei Parameter - tubuläre Formationen, nukleäre Pleomorphismen und Mitosezahl – gemäß den zuvor in der Einleitung beschriebenen Methoden durchgeführt.

Beide Kohorten wurden unter Verwendung unterschiedlicher Mikroskope ausgewertet. Für die erste Kohorte kam das Eclipse CI von Nikon zum Einsatz, während für die zweite Kohorte das Axioskop von Zeiss verwendet wurde. Beide Mikroskope bieten bei einer 40-fachen Vergrößerung einen GFD von 0,50 mm. Die entsprechenden Scores für die Mitosezahl in 10 HPF definieren sich wie folgt: ≤ 7 entspricht Score 1, 8-14 einem Score von 2 und ≥ 15 wird als Score 3 definiert. Die Auswertung konzentrierte sich vor allem auf Bereiche innerhalb der Invasionsfront, sofern im Tumorschnitt nachvollziehbar, oder in der Region mit der höchsten Proliferation.

Eine genaue Zuordnung, welche Strukturen Mitosefiguren entsprechen, kann aus Abbildung 4 entnommen werden.



Abbildung 4: Darstellung der einzelnen Mitosefiguren im HE-Schnitt zugeordnet zu den entsprechenden Teilungsschritten. Mit den roten Pfeilen markiert ist die Sonderform einer tripolaren Mitosefigur, welche ein Zeichen für Tumormalignität ist.

2.3.2 Visuelle Auswertung des Proliferationsindex (Ki67)

Die visuelle Auswertung des Proliferationsindex basierend auf der Ki67-Färbung wurde unter Verwendung von zwei Methoden durchgeführt:

2.3.2.1 Globale Abschätzung

Initial erfolgte eine allgemeine Einschätzung der Proliferationsrate durch eine erfahrene Histologin (Dr. Irene Gonzalez-Menendez). Hierbei wurde in niedriger Vergrößerung (10x-20x) ohne standardisiertes Protokoll eine ungefähre Anzahl der immunmarkierten Zellen abgeschätzt. Die gewonnenen Ergebnisse wurden anschließend in einer Mastertabelle in Excel festgehalten.

2.3.2.2 Globale Methode nach IKWG-Protokoll

Die zweite Methode orientierte sich an dem Protokoll der IKWG (Leung et al., 2019). Zunächst wurde in der gesamten Übersicht (mit maximal 10-facher Vergrößerung) die Proliferationsrate in jedem einzelnen Fall beurteilt. Dabei wurden bestimmte Bereiche identifiziert, die entweder eine hohe, mittlere, niedrige oder negative (d.h. keine Zelle ist gefärbt) Expression von Ki67 aufweisen. Der Anteil dieser Bereiche an der Gesamttumorfläche wurde in Prozent festgelegt.

Anschließend wurde in 40-facher Vergrößerung für jede dieser Regionen ein repräsentatives HPF festgelegt. In diesem HPF wurden die Tumorzellen systematisch von links nach rechts nach einem „Schreibmaschinen-Muster“ (Leung et al., 2019) ausgezählt. Dies erfolgte so lange, bis entweder 100 invasive Zellen oder alle in der Region befindlichen Zellen erfasst waren. Die ermittelten Werte wurden anschließend in die Formel zur Berechnung des gewichteten Scores eingetragen. Diese Berechnungsmethode wird in Abbildung 5 veranschaulicht.

Gewichteter Ki67 score =

$$\frac{\sum_{i \in \{neg, nied., mit., hohe\}} \% \text{ des Schnittes mit } i^{te} \text{ Färbungskategorie} \times \text{totale \# von positiven Tumorkernen gezählt in Feldern mit } i^{te} \text{ Färbungskategorie}}{\text{totale \# von Tumorkernen gezählt in Feldern mit } i^{te} \text{ Färbungskategorie}} \times 100$$

Abbildung 5: Formel zur Berechnung des gewichteten Ki67-Scores gemäß dem Protokoll der IKWG (Leung et al., 2019)

Als positiv werden jene Zellen definiert, bei denen ein nachweislich positiv gefärbter Nukleus erkennbar ist. Die IKWG hat genaue Kriterien für positive Zellkernfärbungen festgelegt. Eine detaillierte Darstellung von positiv gefärbten Nuklei findet sich in Abbildung 6:

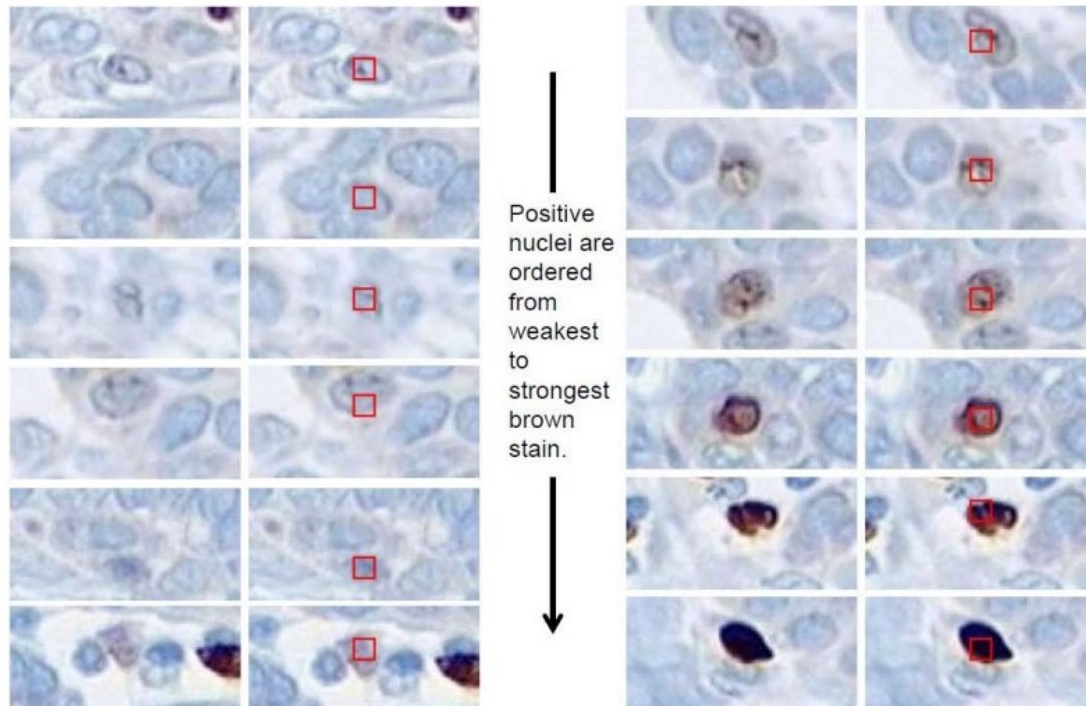
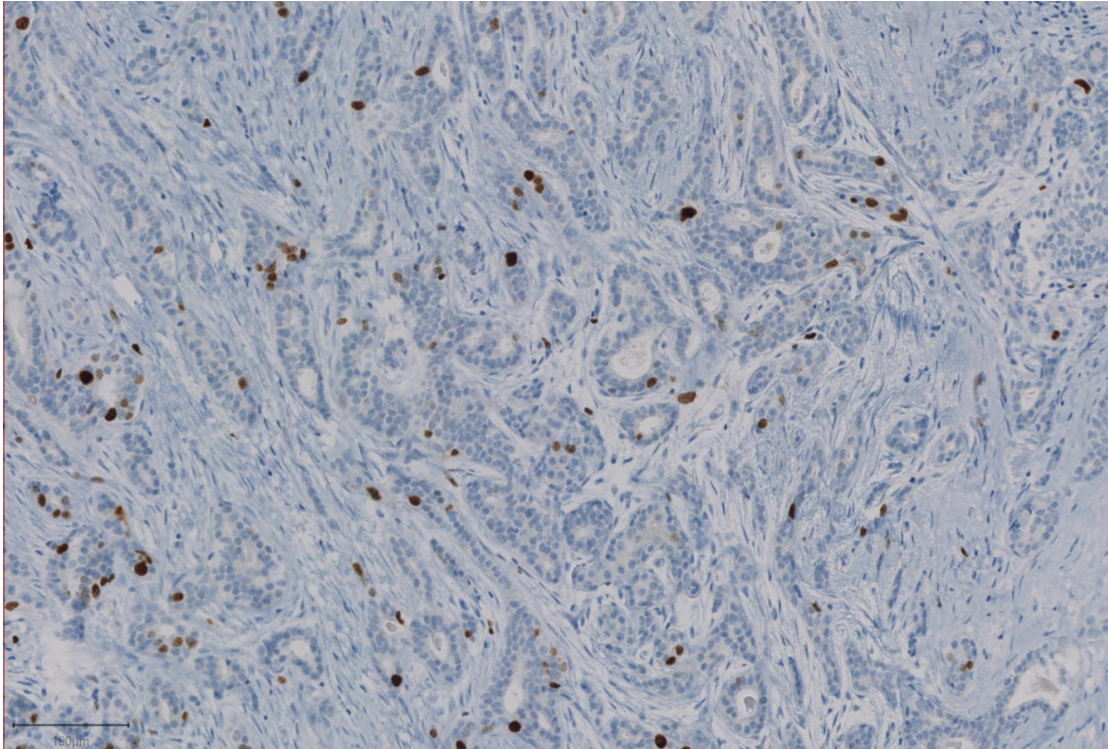
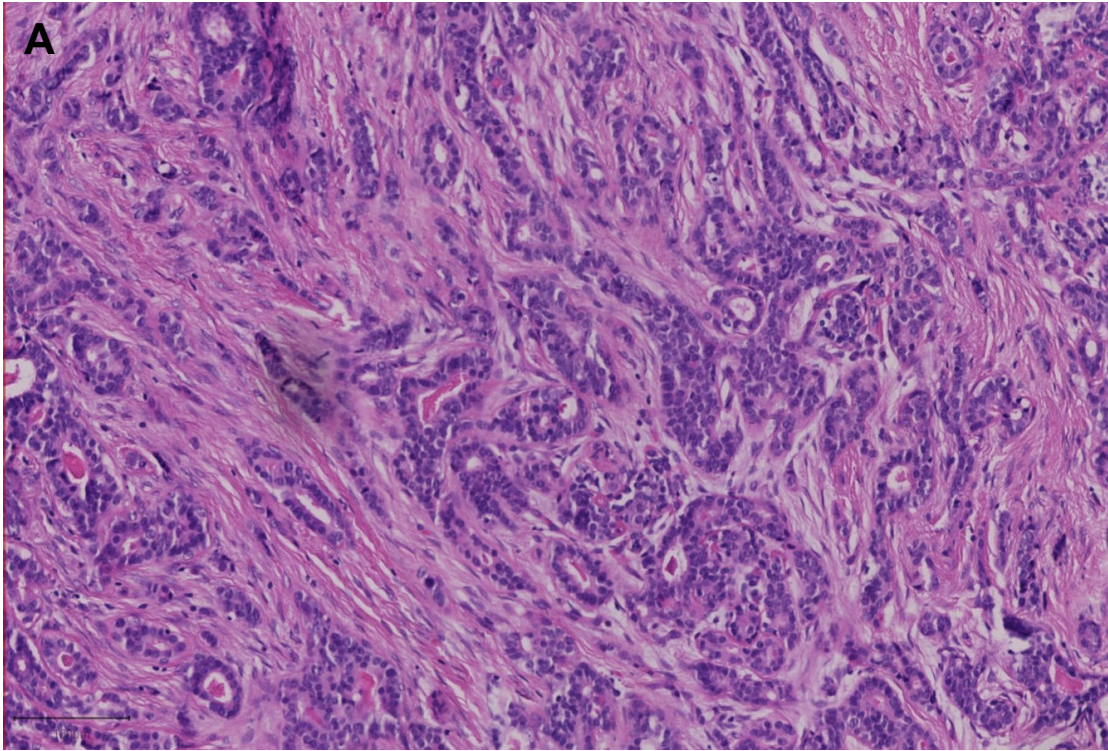
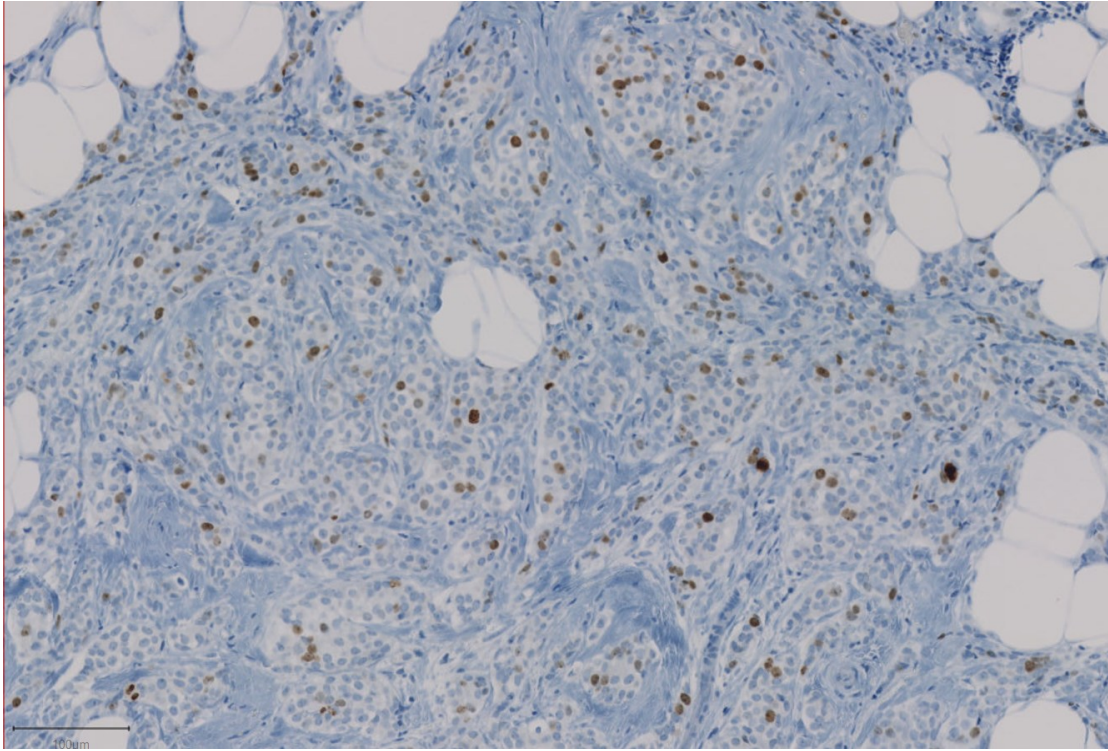
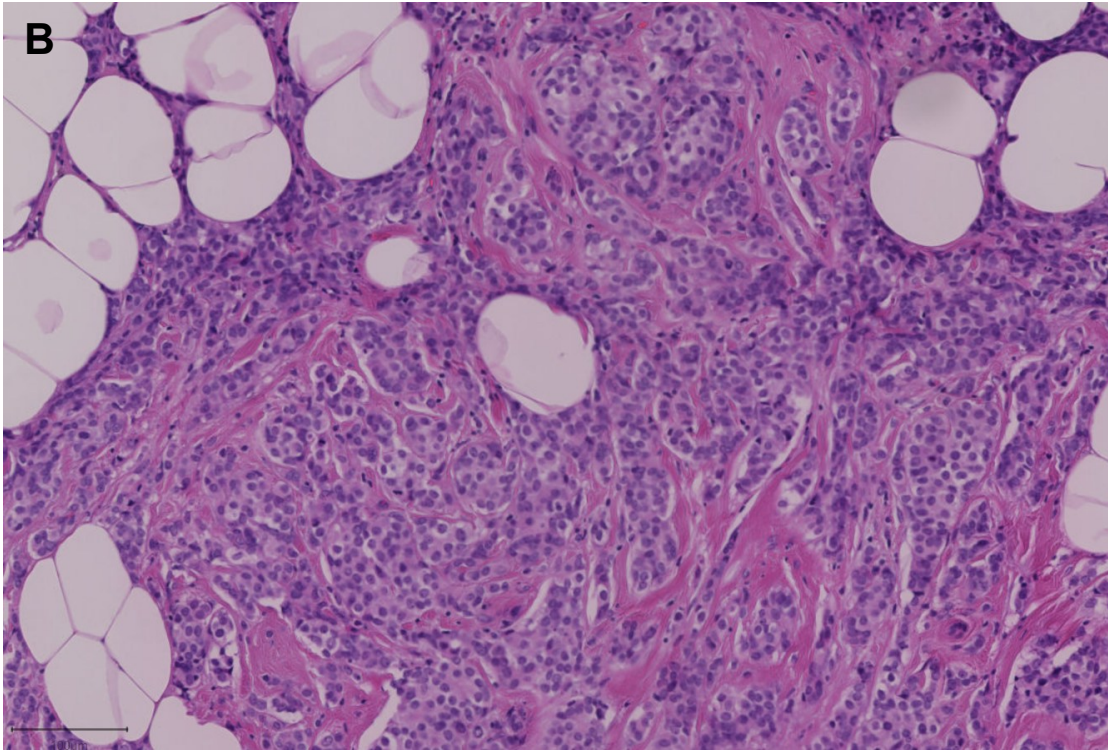


Abbildung 6: Analyse der Ki67-positiven Zellen in Bezug auf ihre Farbintensität gemäß den Empfehlungen der IKWG. Die aufgelisteten Zellen zeigen positive Färbung, von geringster zu stärkster brauner Intensität. Die Färbungsnachweise sind jeweils rot markiert (Leung et al., 2019)





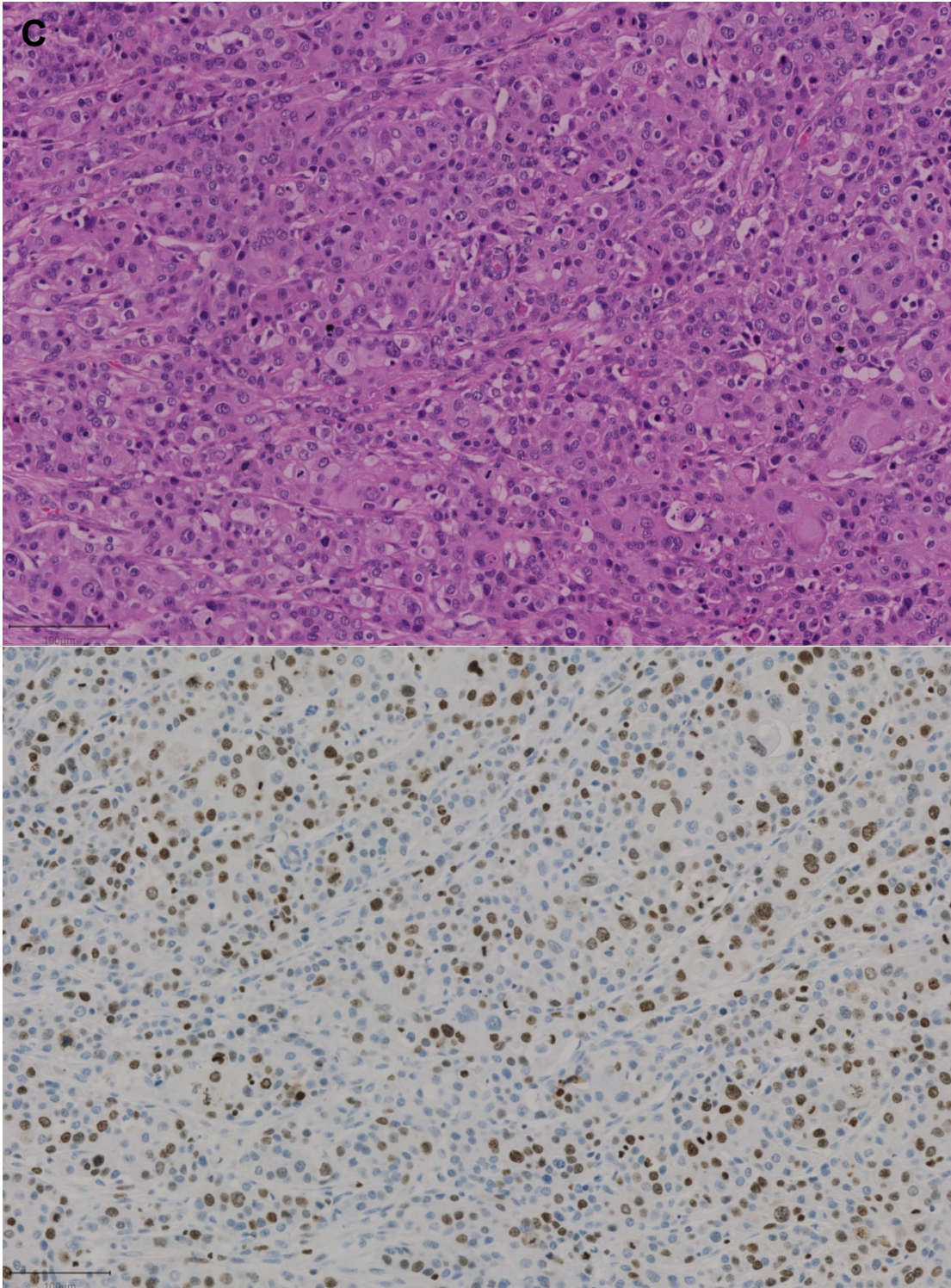


Abbildung 7 Darstellung der Mib1-Färbung bei Fällen mit (A) G1-, (B) G2- und (C) G3-Grading im Vergleich zur HE-Färbung derselben Regionen.

2.3.3 Visuelle Auswertung des Mitoseindex (PhH3)

Die Auswertung der mit PhH3 gefärbten Schnitte der Validierungskohorte erfolgte unter Verwendung eines Axioskop von Zeiss mit einem GFD von 0,50 mm. Die Entdeckungskohorte wurde durch Dr. Ivonne Montes und cand. med. Diana Silimon anhand von digitalen Bildern ausgewertet. Die HPF entsprachen hierbei einem GFD von 0,58 mm. Da es für die Auswertung von PhH3 noch kein standardisiertes Protokoll gibt, diente das Mitosescoring der Nottingham Klassifikation (Galea et al., 1992) als Orientierung. Dieses Scoring folgte, wie bei der Re-Evaluation der HE-Schnitte, den von Ellis definierten Vorgaben für klar definierte GFD (Ellis, 2016). Entsprechend dem größeren GFD für die Auswertung der ersten Kohorte wurde auch der Score angepasst. So entsprachen ≤ 9 Mitosen pro 10 HPF einem Score von 1, 10-19 Mitosen pro 10 HPF erhielten einen Score von 2 und ein Score und ≥ 20 Mitosen pro 10 HPF wurden mit einem Score von 3 bewertet. Um einen Vergleich der Mitosezahlen zwischen der PhH3-Auswertung, der HE-Auswertung und der digitalen Auswertung zu ermöglichen, wurden die ermittelten Mitosezahlen von einer Gesichtsfeldfläche (GFF) von 2,38 mm² auf 1,96 mm² umgerechnet.

Als Mitosen gewertet wurden alle deutlich erkennbaren Mitosefiguren, einschließlich solcher, die falsch-negativ gefärbt waren. Falsch-positiv gefärbte Zellen wurden hingegen ausgeschlossen (siehe Abbildung 8).

Insgesamt wiesen 25 Fälle (19%) nur eine geringe oder keine Reaktion mit PhH3 auf. Betroffen waren ausschließlich Resektate, vor allem mit dem Grading 3 (n=19) und dem molekularen Subtyp TN (n=11) (siehe Tabelle 13). Bei diesen Fällen ließen sich die falsch-negativen Mitosen besonders im Zentrum des Tumors nachweisen. Es ist außerdem erwähnenswert, dass die zugehörigen Stenzen auf dem gleichen Objektträger in allen Fällen eine Färbung aufwiesen und auch falsch-negative Mitosefiguren nicht identifiziert werden konnten.

| | Fallzahl (n) | Anzahl der Fälle (%) |
|-----------------------------|--------------|----------------------|
| Gesamt* | 25 | 19 |
| Histologischer Grad | | |
| 1 | 1 | 0,7 |
| 2 | 5 | 3,8 |
| 3 | 19 | 14,5 |
| Molekularer Subtyp** | | |
| HR+ HER2- (Lum. A) | 4 | 3 |
| HR+ HER2- (Lum. B) | 6 | 4,5 |
| HR+ HER2+ (Lum. B) | 2 | 1,5 |
| HER2+ HR- | 2 | 1,5 |
| TN | 11 | 8,5 |

Tabelle 13: Anzahl der Fälle mit schlechter oder fehlender Reaktion auf PhH3-Färbung. Die größte Gruppe besteht aus Tumoren mit Grad 3 oder einem TN-Subtyp. (*Die Daten beziehen sich nur auf Resektate, ** die Auswertung der Stanzbiopsie bestimmt den klinisch relevanten Subtyp)

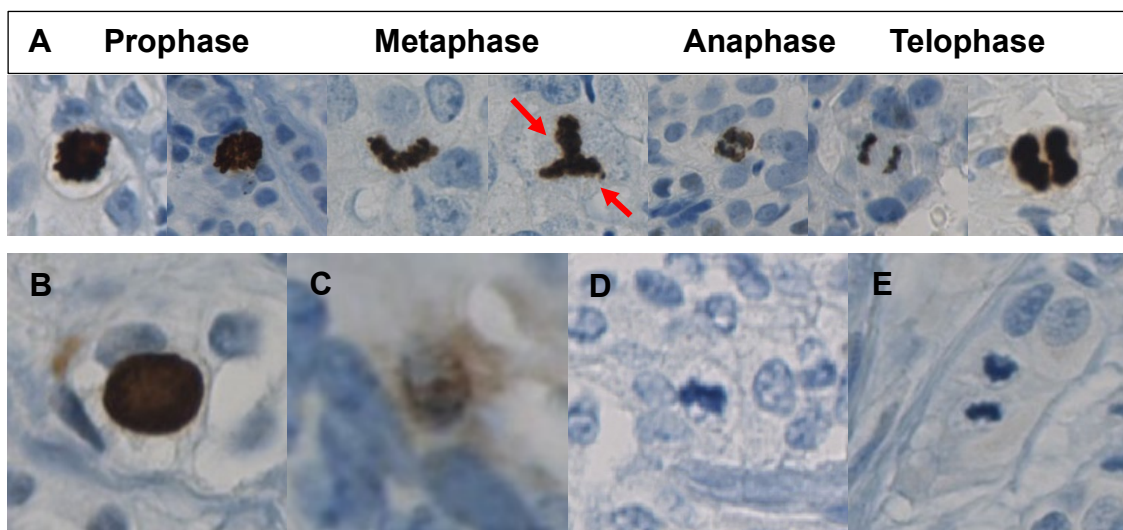
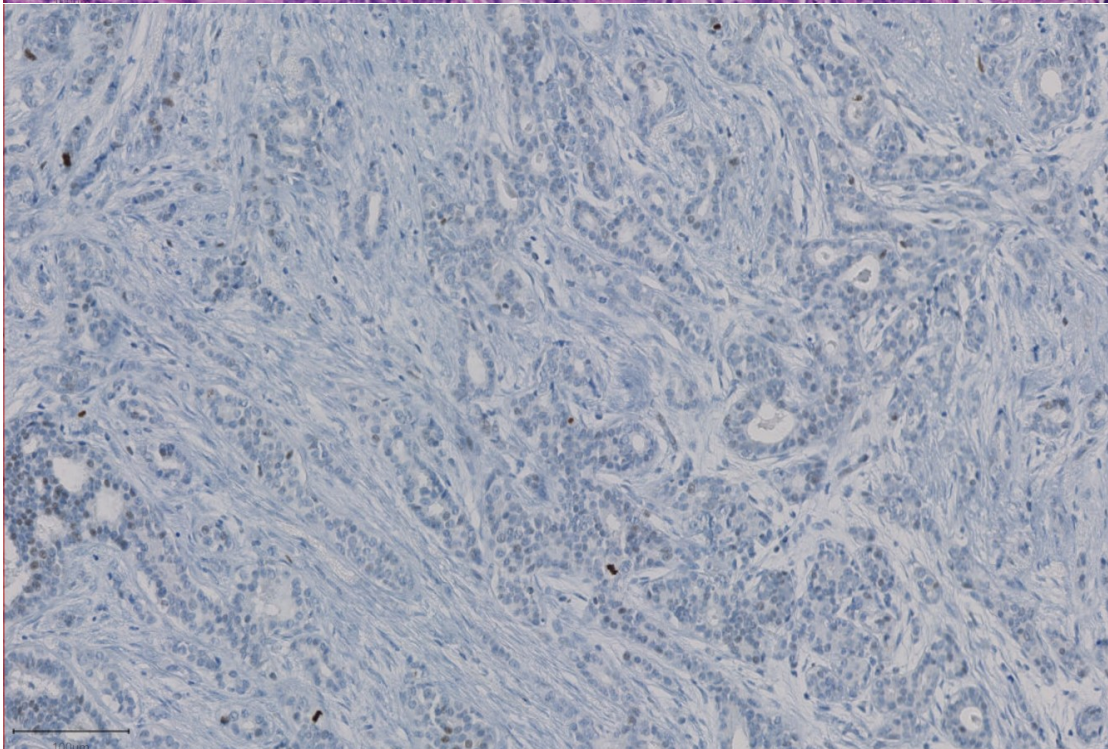
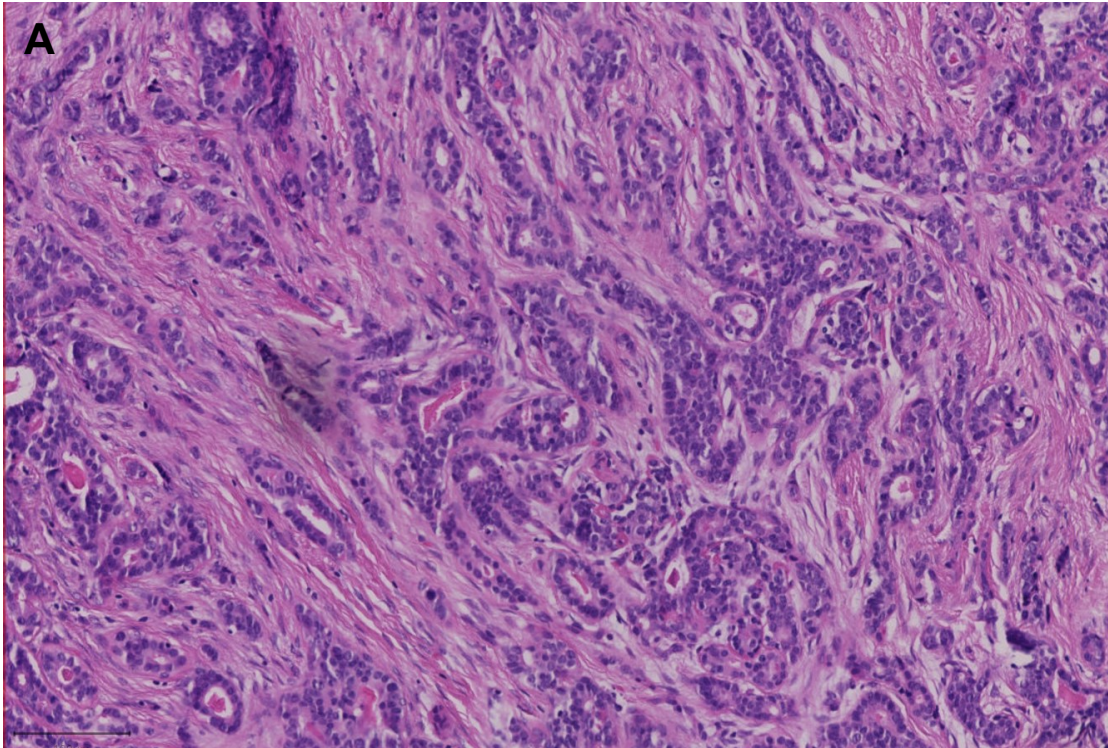
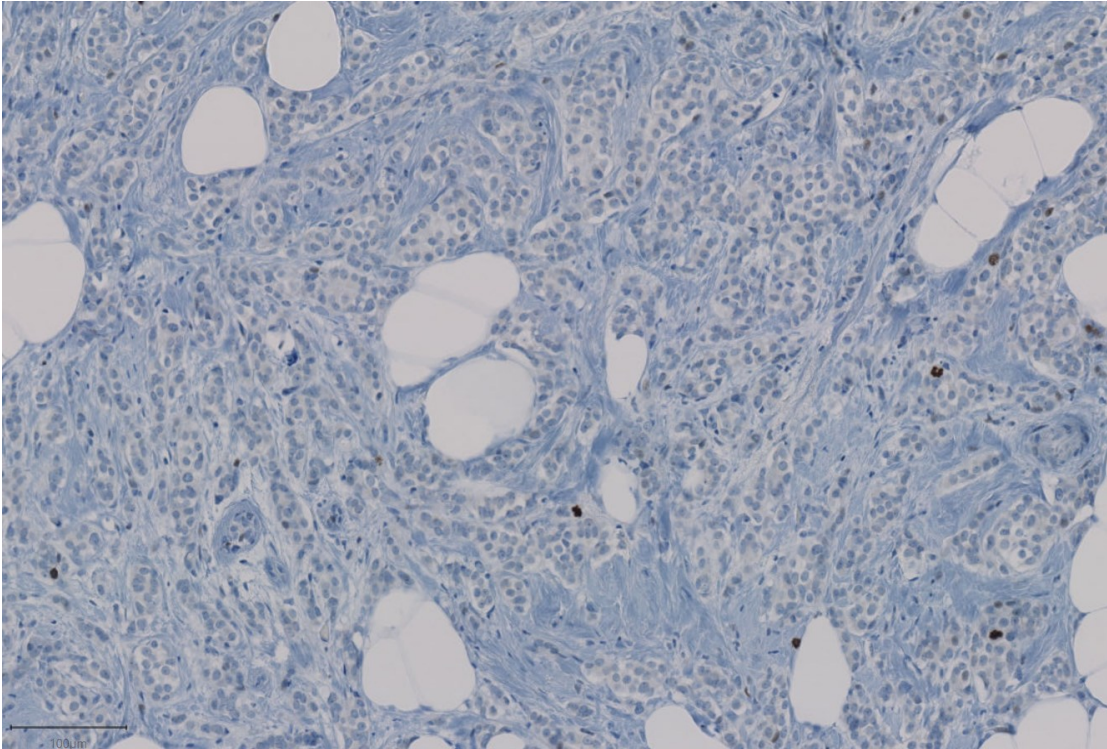
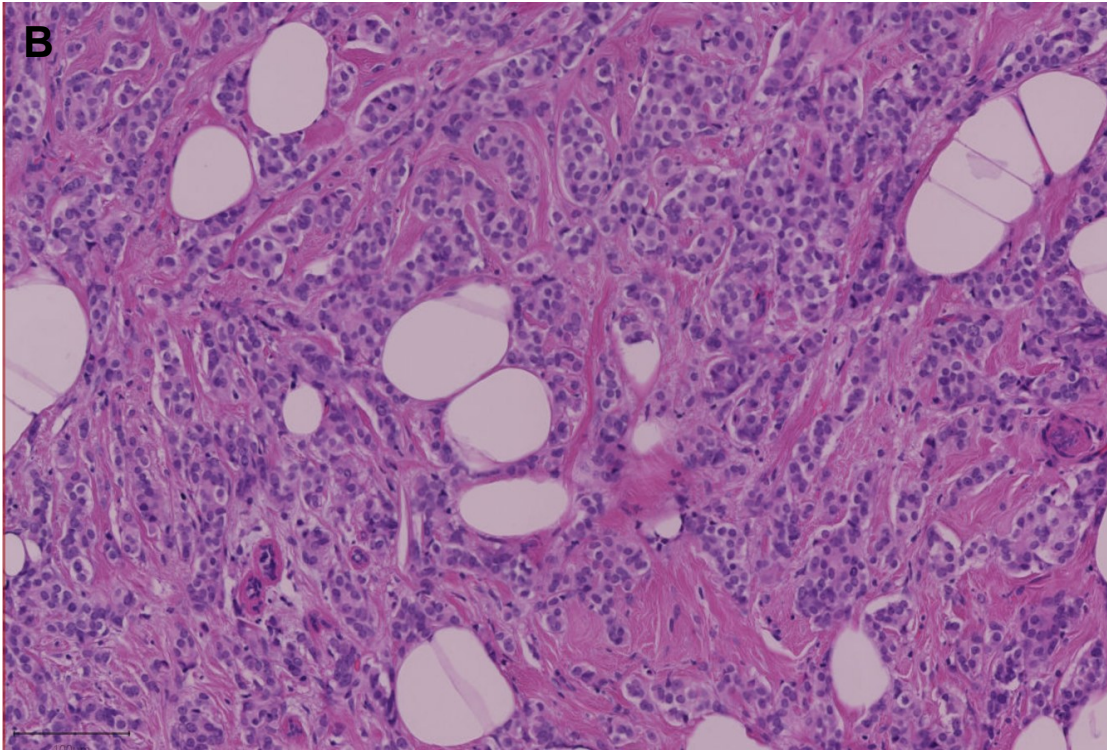


Abbildung 8: Darstellung von Mitosefiguren und Färbeartefakten im PhH3-Schnitt. (A) Zuordnung der einzelnen Mitosefiguren zu den Teilschritten der Zellteilung. Mit den roten Pfeilen markiert ist die Sonderform einer tripolaren Mitosefigur, welche ein Zeichen für Tumormalignität ist. In (B) wird eine stark und in (C) eine schwach falsch positiv gefärbte Zelle dargestellt. (D) Stellt eine sich in der Metaphase und (E) in der Telophase befindliche falsch negativ gefärbte Mitosefigur dar.





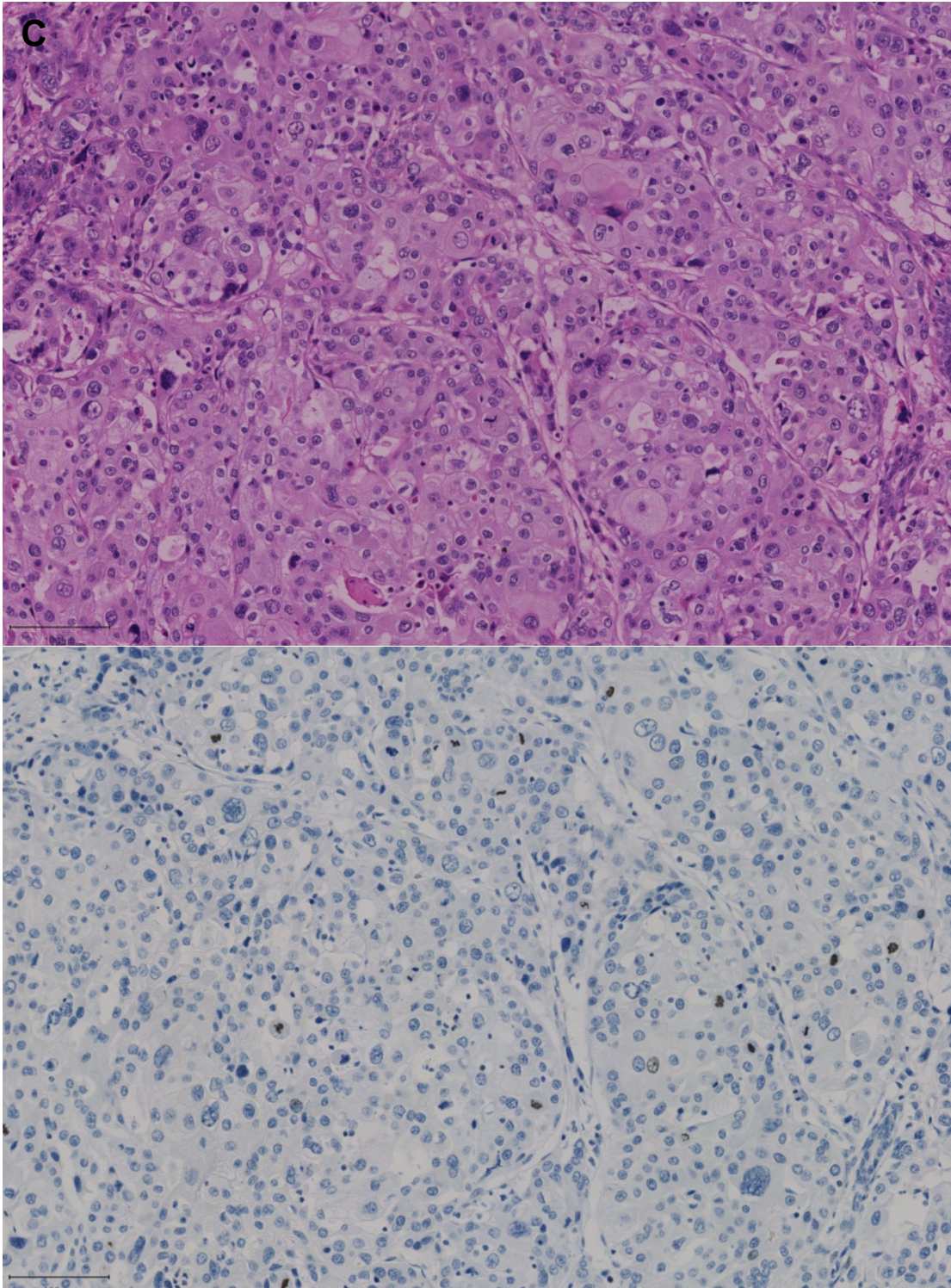


Abbildung 9 Darstellung der PhH3-Färbung bei Fällen mit (A) G1-, (B) G2- und (C) G3-Grading im Vergleich zur HE-Färbung derselben Regionen.

2.4 Digitale Auswertung

Die Digitalisierung der histologischen Schnitte erfolgte für die erste Kohorte mithilfe des Ventana DP 200-Scanners von Roche im Bildformat .tif. Die zweite Kohorte wurde mit einem NanoZoomer 2.0 HT von Hamamatsu erfasst, der Bilder im .ndpi-Format generiert. Die anschließende digitale Auswertung erfolgte unter Verwendung der Open-Source-Software QuPath, welche eine große Bandbreite von Bildformaten unterstützt und somit eine universelle Anwendung ermöglicht. Alle Bilder wurden unter den entsprechenden K-Nummer (Fallnummer im digitalen System der Pathologie – Nexus) auf einer externen Festplatte gespeichert (Seagate Expansion Desktop Festplatte 8TB HDD).

Um eine strukturierte Einordnung der Fälle zu gewährleisten, wurde eine Ordnerstruktur etabliert. Dabei wurde jedem Fall eine spezifische Studiennummer zugeordnet, die aus verschiedenen Komponenten besteht. Diese beinhaltet die Kohortenzugehörigkeit (1 = Entdeckungskohorte, 2 = Validierungskohorte), das Grading bzw. den molekularen Subtyp (1 = Grad 1 HR+, 2 = Grad 2 HR+, 3 = Grad 3 HR+, 4 = HER2+, 5 = TN) und die individuelle Fallnummer innerhalb des Subtyps (23 = 23. Fall der Untergruppe). Ein exemplarischer Fall mit der Nummer 22023 würde also dem 23. HR+ Fall mit einem Grading von 2 aus der Validierungskohorte entsprechen.

Innerhalb der QuPath-Software wurden insgesamt vier separate Projekte angelegt. Diese waren nach Stanzbiopsien und Resektaten sowie nach der Auswertung von Ki67 und PhH3 strukturiert. Die digitalisierten Bilder wurden gemäß der Reihenfolge der zugeordneten Studiennummer in das jeweilige Projekt integriert.

2.4.1 Digitale Auswertung des Proliferationsindex (Ki67)

Die digitale Auswertung des Ki67-Proliferationsindex basierte auf dem Protokoll der IKWG (Acs et al., 2019). Aufgrund der fortlaufenden Weiterentwicklung von QuPath waren leichte Anpassungen im Protokoll notwendig, insbesondere in Bezug auf das Training eines „Classifiers“ (s.u.).

Nach dem erfolgreichen Hochladen der Bilder wurde der erste Schritt der digitalen Auswertung durchgeführt: Die Umrandung der Tumorfläche mithilfe des Annotationstools. Dabei wurde darauf geachtet, dass große Flächen Stroma, Nekrosen oder sonstige Artefakte nicht eingeschlossen wurden. Anschließend wurde mit dem Rechteck-Tool ein kleiner Bereich festgelegt, der möglichst verschiedene Färbintensitäten repräsentierte. Unter dem Reiter „Analyze/Preprocessing/Estimate stain vectors“ erfolgte die automatische Berechnung der Färbungsintensität. Um die Werte an das aktuelle Bild anzupassen, wurde im angezeigten „Visual Stain Editor“ die Schaltfläche „Auto“ ausgewählt.

Im nächsten Schritt wurde erneut die umrandete Tumorfläche ausgewählt. Unter „Analyze/Cell analysis/Positive cell detection“ öffnete sich ein neues Fenster. Hierbei wurde bei der Option „Choose detection image“ die Einstellung „Optical density sum“ ausgewählt. Die genauen Einstellungen für diesen Schritt können aus Abbildung 10 entnommen werden.

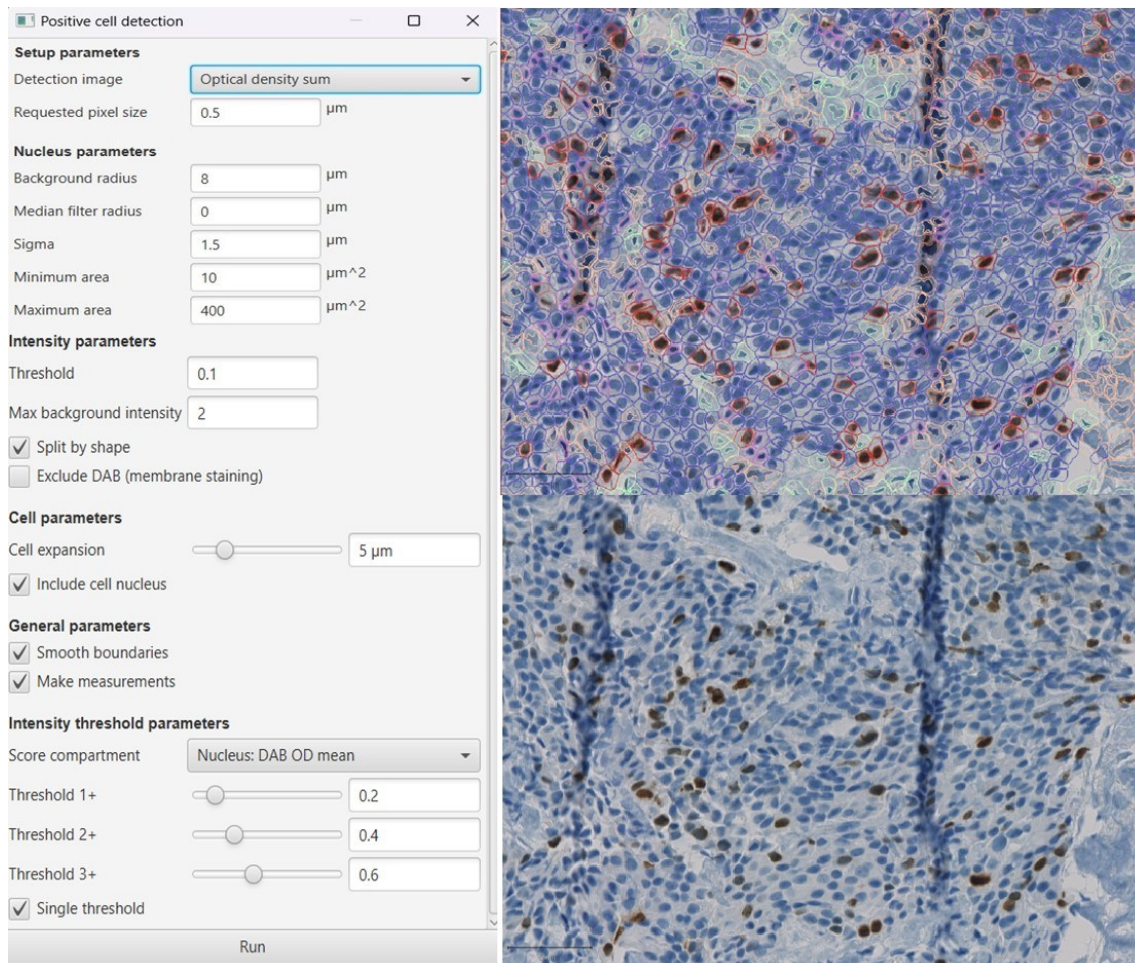


Abbildung 10: Zeldetektionstool der Software QuPath zur Auswertung de Ki67-Proliferationsindex. Links: Darstellung der genauen QuPath-Einstellung für die „Positive cell detection“. Rechts: Digitale Darstellung von Tumorgewebe mit und ohne Zeldetektion. Rote Farbe markiert positive Zellen, Grün repräsentiert das Stroma und Orange kennzeichnet Bildartefakte.

Um eine noch präzisere Klassifikation zu ermöglichen, bietet QuPath die Möglichkeit, zusätzliche Eigenschaften zu bestimmen. Hierzu wurde die Funktion „Analyze/Calculate features/Add smoothed feature“ verwendet, um individuelle Zellmaße von 25 μm und 50μm zu erfassen.

Zum Training der KI von QuPath wurde im Anschluss ein Klassifikator erstellt. Dieser konnte über „Classify/Object classification/Train object classifier“ aufgerufen und mit der Option “Live update” aktiviert werden. Informationen aus vorherigen Bildanalysen konnten unter „Load training“ hinzugefügt werden. Durch die Verwendung des „Polygon“-Tools war es möglich, Bereiche zu markieren, die vom Klassifikator fehlerhaft identifiziert wurden. Mit einem Rechtsklick auf den markierten Bereich unter dem Reiter „Set class“ konnte man

die Detektion neu klassifizieren. In unserem Klassifikator wurde hierfür unter anderem die Klasse „Artefakt“ hinzugefügt. Diese steht für Auffaltungen des Präparates, Schmutzpartikel, fehlerhafte Antikörperreaktionen etc.

Wurde nach der Kontrolle und einer eventuellen Korrektur durch die analysierende Person ein zufriedenstellendes Ergebnis festgestellt, so konnte dieses im Datenfenster unten links unter dem Punkt „Positive %“ eingesehen und in die Mastertabelle übertragen werden.

2.4.1.1 VM-Scope (Hot Spot-Methode)

Zur ergänzenden digitalen Auswertung der ersten Kohorte wurde die kommerzielle Bildanalyse-Software „Cognition Master Professional Suite“ der VM Scope GmbH eingesetzt. Dabei kam das Mikroskop Zeiss Axio Imager.A1 zum Einsatz. Dieses ermöglicht die simultane Aufnahme von Bildern, die von der Software in Echtzeit analysiert werden können.

In jedem Schnitt wurden 10 Hot Spots identifiziert und mithilfe des „Ki67 Quantifyer“ die durchschnittliche Proliferationsrate ermittelt. Dieses Tool erkennt automatisch sowohl negative als auch positive Tumorzellen und kann diese gleichzeitig vom umliegenden Stroma differenzieren, wobei sich die Auswertung ausschließlich auf den invasiven Tumor konzentrierte. Die individuellen Proliferationsraten wurden automatisch zu einem Durchschnittswert verrechnet, der anschließend für die statistische Auswertung in die Mastertabelle übertragen wurde.

2.4.2 Digitale Auswertung des Mitoseindex (PhH3)

Die digitale Auswertung von PhH3 ist ebenso wie die visuelle Auswertung bislang nicht standardisiert. Daher wurde für diese Studie das bereits erwähnte Protokoll zur Ki67-Auswertung modifiziert. Hierbei wurde das Ziel verfolgt, eine möglichst präzise und reproduzierbare Methode zur globalen Bestimmung des Mitoseindex zu entwickeln. Dabei sollte die exakte Anzahl von Mitosen auf der Tumorfläche berechnet werden. Zu diesem Zweck wurden die einzelnen Tumornester so genau wie möglich mit den Annotationstools markiert, um einen Einfluss von Stroma, Artefakten oder Nekrosen auf die Gesamtfläche so gering wie möglich

zu halten. Im Anschluss wurden die markierten Annotationen mit der Funktion „Select all/Rechtsklick/Edit multiple/Merge selected“ zusammengefügt, um die Gesamtfläche in μm^2 im Datenfenster unter dem Reiter „Area μm^2 “ abzulesen.

Die folgenden Schritte entsprechen weitestgehend denen der Ki67-Auswertung, darunter die Bestimmung der „Stain vectors“ und die „Positive cell detection“. Die für die Mitosebestimmung spezifischen Einstellungen der Zelldetektion können aus Abbildung 11 entnommen werden.

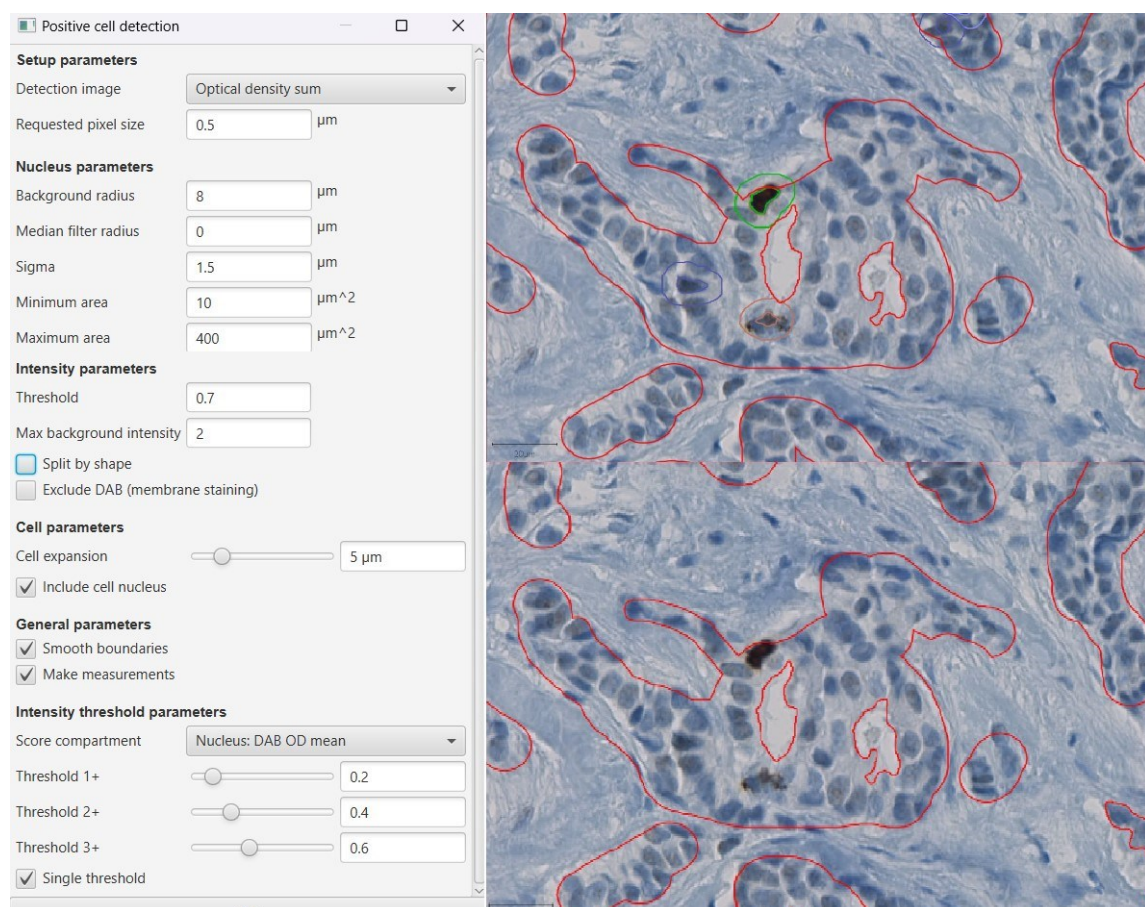


Abbildung 11: Zelldetektionstool der Software QuPath zur Auswertung der Mitoserate anhand von PhH3-gefärbten Schnitten. Darstellung der genauen QuPath-Einstellung für die „Positive cell detection“. Rechts: Digitale Darstellung von Tumorgewebe mit und ohne Zelldetektion. Grüne Farbe markiert Mitosen, Orange kennzeichnet Bildartefakte. Rote Annotationen zeigen die präzise Markierung der einzelnen Tumornester.

Im anschließenden Schritt werden die „smoothed features“ für die markierte Tumorfläche bestimmt und der Klassifikator trainiert (s.o.). Die erzielten Ergebnisse wurden im Anschluss ebenfalls nach Bedarf korrigiert, bis sie für die analysierende Person zufriedenstellend waren. Für die Analyse der Mitosen

wurde darüber hinaus die Klasse „Mitosis“ den entsprechenden QuPath-Projekten hinzugefügt.

Die Anzahl der identifizierten Mitosen wurde mit der annotierten Gewebefläche in Beziehung gesetzt und im Anschluss auf eine Flächengröße von 1,96 mm² umgerechnet, damit es der GFF des verwendeten Mikroskops entsprach.

2.5 Statistische Auswertung

Die statistische Auswertung erfolgte in dieser Studie mithilfe der Softwareprogramme SPSS 29 (Statistical Package for the Social Sciences) von IBM und R, einer Open-Source-Programmiersprache, welche von Statistikern der Universität Auckland entwickelt wurde. Die Auswertung wurde mit Unterstützung des Instituts für klinische Epidemiologie und angewandte Biometrie (IKEaB) Tübingen (Lina Maria Serna Higueta) und der Pathologin Dr. Veronika Bahlinger durchgeführt.

Die Ergebnisse wurden zur Veranschaulichung in verschiedenen Diagrammformen dargestellt, darunter Sankey-Plots, Boxplots, Punktdiagramme, Balkendiagramme und gestapelte Balkendiagramme.

Des Weiteren wurden Clusteranalysen durchgeführt, um Patientinnen mit ähnlichen Merkmalen zu gruppieren. Hierbei wurden sogenannte K-Medoid-Cluster erstellt, wobei als Grundlage für die Clusteranalysen vier Parameter festgelegt wurden: die tubuläre Formation, nukleärer Pleomorphismus, visuelle Auswertung von PhH3 und die digitale Auswertung von Ki67 mithilfe von QuPath.

Die Analyse des Gesamtüberlebens und des krankheitsfreien Überlebens in Relation zu den einzelnen Clustern und dem Grading der Validierungskohorte wurde außerdem in Form von Kaplan-Meyer-Kurven dargestellt. Zum Vergleich der Unterschiede zwischen den Gruppen wurde der Log-Rank-Test verwendet. Die Follow-Up-Zeiträume erstreckten sich bis zu 140 Monate nach dem Diagnosezeitpunkt oder bis zum Auftreten eines Lokalrezidivs, einer Lymphknoten- oder Fernmetastase (Rezidivfreies Überleben) oder bis zum Tod der Patientin (Gesamtüberleben). Es ist zu beachten, dass nach Bereitstellung

der klinischen Daten 10 Fälle (nach Re-Grading: G1 HR+ HER2-: n=1; G2 HR+ HER2-: n=2; G3 HR+ HER2- n=4; G3 HR+ HER2+: n=1; G3 TN: n=2) aufgrund von vorangegangenen Tumordiagnosen für die Analyse der Überlebensdaten ausgeschlossen werden mussten (Mammakarzinom: n=7, Schilddrüsenkarzinom: n=1, Hodkin-Lymphom: n=1, Cervixkarzinom n=1).

Des Weiteren entsprechen die gesammelten Daten nicht der Normalverteilung, was durch den Shapiro-Wilk-Test bestätigt wurde. Nichtsdestotrotz, aufgrund der ausreichenden Stichprobengröße von mehr als 30 Proben, ist eine Normalverteilung für die Korrelation der Daten nicht erforderlich. Zur Untersuchung von Zusammenhängen zwischen den verwendeten Methoden und Probenformen wurde die Pearson-Korrelation herangezogen. Da sämtliche statistischen Tests zweiseitig sind, wurden p-Werte von weniger als 0,05 als statistisch signifikant angesehen. Bei der Korrelationsanalyse galten R-Werte größer als 0,3 als moderate Korrelation, während R-Werte größer als 0,5 als starke Korrelation betrachtet wurden.

3 Ergebnisse

3.1 Entdeckungskohorte

3.1.1 Re-Grading

Durch das Re-Grading der einzelnen Fälle anhand des HE-Schnittes ergaben sich Veränderungen in der Graduierung bei insgesamt fünf Fällen. Während die Fälle, die zuvor als Grad 1 eingestuft wurden, ihr ursprüngliches Grading beibehielten, wurde bei vier Fällen mit anfänglichem Grad 2 eine Hochstufung zu Grad 3 vorgenommen. Lediglich ein Fall wurde von Grad 3 zu Grad 2 heruntergestuft. Die größte Gruppe setzt sich sowohl vor als auch nach Re-Grading aus den Fällen mit Grad 3 zusammen (nach Re-Grading, n=47, 65,3%). Die Fälle mit Grad 1 und Grad 2 sind nahezu gleichmäßig verteilt (G1: n=12, 16,7%, G2: n=13, 18,1%).

Bei einer erneuten Einschätzung der Fälle mittels PhH3-Mitosezählung (orientiert an den Vorgaben des Gradings) sei es visuell oder mithilfe der digitalen Methode (QuPath), ergab sich eine umfassende Neuklassifikation der meisten Fälle. Lediglich 2,8% der Fälle (n=2) wurden noch als Grad 1 und 12,5% (n=9) als Grad 2 eingestuft. Ähnliche Ergebnisse zeigten sich bei Analyse durch die digitale Methode. Eine detaillierte Darstellung der Verteilung der Fälle kann aus Abbildung 12 und Tabelle 14 entnommen werden.

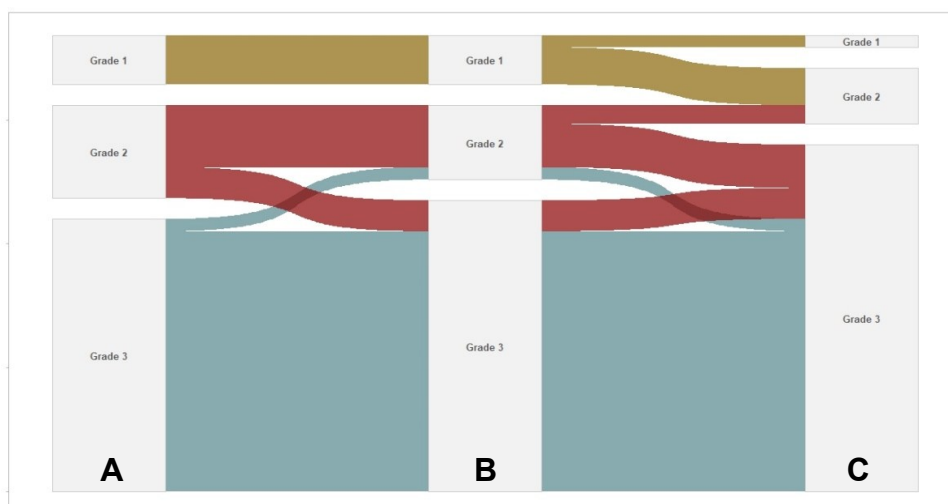


Abbildung 12: Sankey-Plot zur Darstellung der Veränderungen des Gradings nach den einzelnen Auswertungsmethoden (Entdeckungskohorte). Dabei werden (A) die Original-Gradings, (B) die Re-Evaluation (HE) und (C) die Re-Klassifikation nach visueller PhH3-Auswertung aufgeführt. Nach der visuellen Auswertung mit PhH3 wird der größte Anteil der Fälle als G3 eingestuft.

| | Gesamt (N=72) |
|--|----------------------|
| Grading Originalbefund | |
| 1 | 12 (16.7%) |
| 2 | 16 (22.2%) |
| 3 | 44 (61.1%) |
| Re-Evaluation des Gradings | |
| 1 | 12 (16.7%) |
| 2 | 13 (18.1%) |
| 3 | 47 (65.3%) |
| Re-Klassifikation nach visuellem Mitosescore (pHH3) | |
| 1 | 2 (2.8%) |
| 2 | 9 (12.5%) |
| 3 | 56 (77.8%) |
| Fehlend | 5 (6.9%) |
| Re-Klassifikation mit QuPath Mitosescore | |
| 1 | 5 (6.9%) |
| 2 | 11 (15.3%) |
| 3 | 56 (77.8%) |

Tabelle 14: Re-Klassifikation nach Re-Evaluation und PhH3-Auswertung (Entdeckungskohorte). Die Tabelle zeigt die Anpassungen der Klassifikationen nach den einzelnen Re-Evaluationsmethoden (Entdeckungskohorte). Auffällig hierbei ist die hohe Anzahl der Fälle, die nach der visuellen und digitalen PhH3 Analyse als Grad 3 klassifiziert wurden.

3.1.2 Visuelle und digitale Auswertung des Ki67-Proliferationsindex

Der Vergleich der angewandten visuellen und digitalen Analysemethoden innerhalb dieser Kohorte verdeutlicht eine bemerkenswerte Ähnlichkeit der erzielten Resultate. Bei insgesamt 64 ausgewerteten Fällen zeigt sich bei den visuellen Methoden lediglich eine geringfügige Divergenz in den Medianen (M) sowie in den Interquartilabständen (IQ) (Visuelle Schätzung M: 35%, IQ: 37,5%; IKWG Protokoll M: 37%, IQ: 39,35%). Ein ähnliches Muster zeigt bei der Auswertung durch QuPath (M: 38,62%, IQ: 42,88%). Dies unterstreicht, dass die Diskrepanz zwischen allen vier angewandten Methoden eher gering ausfällt. Lediglich die Hot-Spot-Auswertung mittels VM-Scope fällt im Median um etwa 5% höher aus (VM-Scope M: 43%, IQ: 36,1%) (Abbildung 13).

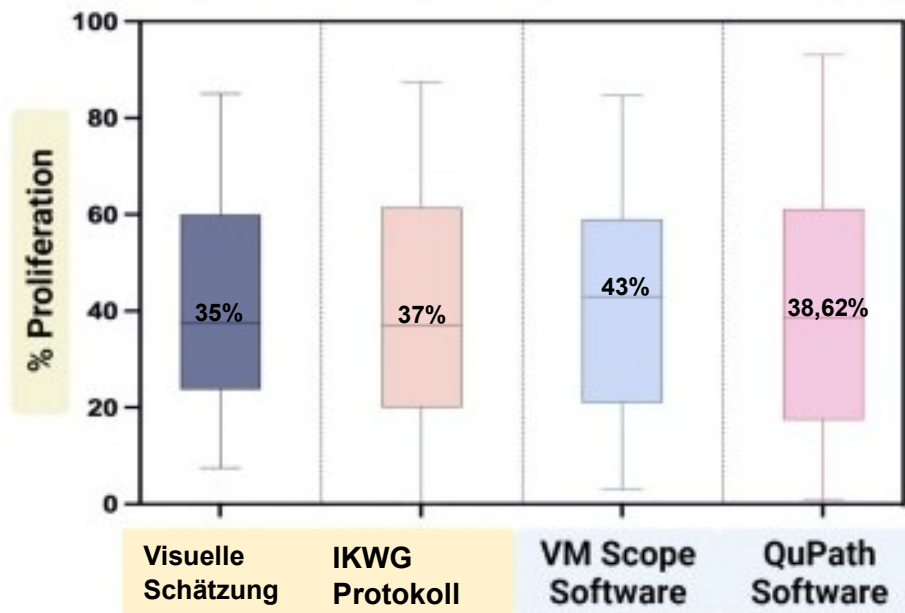


Abbildung 13: Boxplots zur Darstellung der Proliferationsindizes durch visuelle und digitale Auswertung (Entdeckungskohorte). Die visuellen Methoden sind gelb markiert, während die digitalen Methoden blau markiert sind. Die Ergebnisse zeigen nur geringfügige Unterschiede zwischen den beiden Methoden. Diagramm basierend auf Montes et al. (I. A. Montes Mojarro and Fend, 2023)

Bei der detaillierten Analyse der einzelnen Methoden wird deutlich, dass die visuellen Analysemethoden eine höhere Anzahl von Fällen mit einer Proliferationsrate von >20% identifizieren (siehe Abbildung 14). Genauer wurden durch das IKWG-Protokoll 31,25%, Schätzung 29,7%, QuPath 20,3% und VM-Scope 21,9% in diese Kategorie eingestuft. Die Implikation dieser Ergebnisse für die Klassifizierung luminaler Tumoren werden im nachfolgenden Abschnitt näher erörtert (siehe Kapitel 3.1.2.1). Ebenfalls konnte festgestellt werden, dass die angewendeten digitalen Methoden eine höhere Anzahl von Fällen mit einer Proliferationsrate >40% erfassen (IKWG: 46,9%; Schätzung: 46,9%; QuPath: 54,7%; VM-Scope: 56,25%).

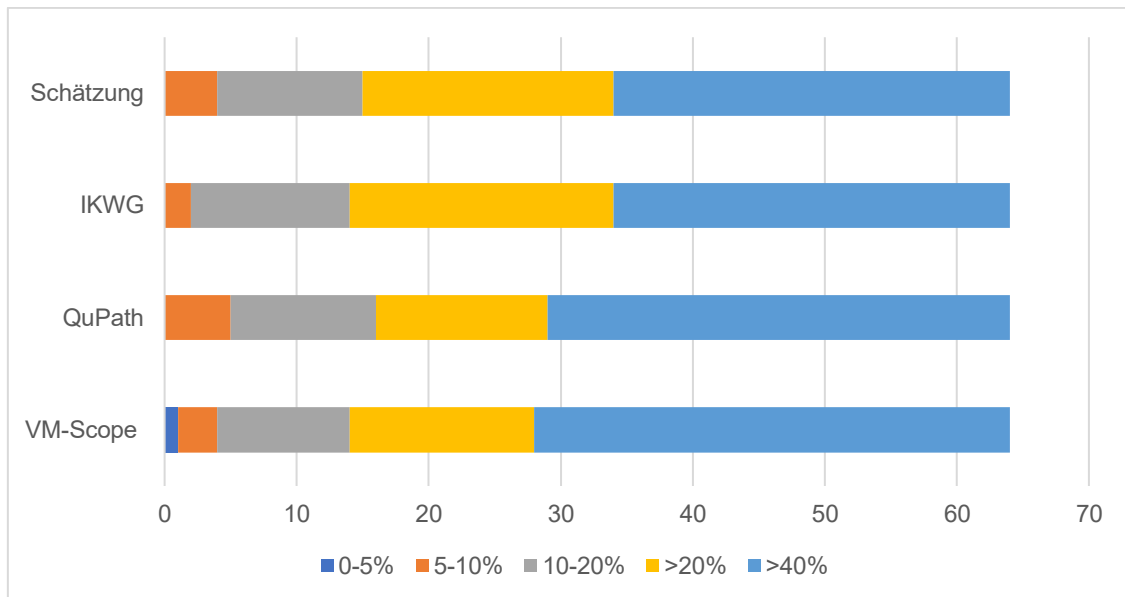


Abbildung 14: Gestapeltes Balkendiagramm mit der Aufteilung der visuellen und digitalen Methoden nach Gruppen unterschiedlicher Proliferationsraten in % (Entdeckungskohorte). Die Gruppe der Fälle mit einer Proliferation von >20% ist bei den visuellen Methoden (Schätzung, IKWG-Protokoll) etwas größer als bei den digitalen Methoden (QuPath, VM-Scope). Hingegen weisen letztere einen höheren Anteil an Fällen über 40% Proliferationsaktivität auf.

Im Rahmen der Korrelationsanalyse nach Pearson der einzelnen Fälle wurde festgestellt, dass alle angewandten Methoden signifikant miteinander korrelieren ($p < 0,001$, $R: 0,719-0,985$). Am stärksten korrelieren hierbei die beiden visuellen Analysemethoden miteinander ($p < 0,001$, $r: 0,985$).

Die digitalen Analysemethoden zeigen ebenfalls eine signifikante Kohärenz, wenn auch mit einer etwas geringeren Korrelation ($p < 0,001$, $R: 0,765$). Beim Vergleich der digitalen und visuellen Analysemethoden wird ersichtlich, dass die visuelle Schätzung und das IKWG-Protokoll jeweils eine gleichwertige Korrelation mit der QuPath-Methode aufweisen ($p < 0,001$, $R: 0,884$).

Die herausgearbeiteten Zusammenhänge zwischen den einzelnen Methoden und deren Beziehung zur standardisierten IKWG-Methode, sowie die Korrelationsstärken untereinander, werden in detaillierter Form in Abbildung 15 visualisiert.

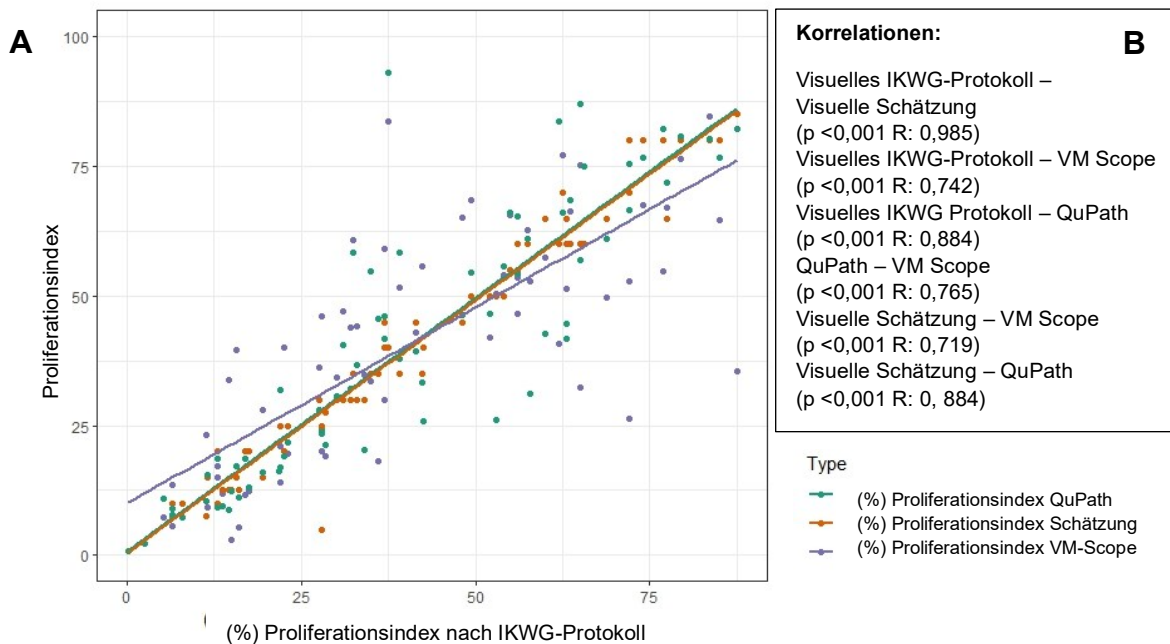


Abbildung 15: Darstellung der Korrelationen nach Pearson von digitalen und visuellen Methoden zur Auswertung des Ki67-Proliferationsindex (Entdeckungskohorte). (A) Der Scatterplot zeigt eine hohe Korrelation der angewendeten Analysemethoden von Ki67 mit der Auswertung des Proliferationsindex nach dem IKWG-Protokoll. Auf der rechten Seite (B) sind die einzelnen Signifikanzniveaus und Korrelationskoeffizienten zusätzlich dargestellt.

3.1.2.1 Veränderung des molekularen Subtyps nach Re-Evaluation des Ki67-Proliferationsindex

Der Ki67-Proliferationsindex nimmt eine signifikante Rolle im Hinblick auf den klinischen Verlauf und die therapeutischen Aspekte von luminalen Tumoren ein. In dieser Studie zeigte sich bei der visuellen Auswertung mittels Schätzung eine vermehrte Kategorisierung von Fällen als Luminal B (Luminal A: 12,5%; Luminal B: 20,3%). Ähnlich verhält es sich bei der Auswertung durch VM-Scope (Luminal A: 10,9%; Luminal B: 21,8%).

Im Zuge des Vergleichs zwischen den globalen Auswertungsmethoden (visuelle Analyse nach dem IKWG-Protokoll und digitale Analyse mittels QuPath) wird außerdem erkennbar, dass in der visuellen Bewertung eine höhere Anzahl von Fällen als Luminal B eingestuft wird (20,3%) im Gegensatz zur digitalen Auswertung (17,1%) (siehe Tabelle 15, Abbildung 16)

| | Gesamt (N=64) |
|---|----------------------|
| Molekularer Subtyp Originalbefund | |
| HR+ HER2- (Lum. A) | 10 (15,6%) |
| HR+ HER2- (Lum. B) | 11 (17,1%) |
| HR+ HER2+ (Lum. B) | 18 (28,1%) |
| HER2+ HR- | 11 (17,2%) |
| TN | 14 (21,9%) |
| Molekularer Subtyp nach Schätzung | |
| HR+ HER2- (Lum. A) | 8 (12,5%) |
| HR+ HER2- (Lum. B) | 13 (20,3%) |
| HR+ HER2+ (Lum. B) | 18 (28,1%) |
| HER2+ HR- | 11 (17,2%) |
| TN | 14 (21,9%) |
| Molekularer Subtyp nach IKWG-Protokoll | |
| HR+ HER2- (Lum. A) | 8 (12,5%) |
| HR+ HER2- (Lum. B) | 13 (20,3%) |
| HR+ HER2+ (Lum. B) | 18 (28,1%) |
| HER2+ HR- | 11 (17,2%) |
| TN | 14 (21,9%) |
| Molekularer Subtyp nach VM-Scope | |
| HR+ HER2- (Lum. A) | 7 (10,9%) |
| HR+ HER2- (Lum. B) | 14 (21,8%) |
| HR+ HER2+ (Lum. B) | 18 (28,1%) |
| HER2+ HR- | 11 (17,2%) |
| TN | 14 (21,9%) |
| Molekularer Subtyp nach QuPath | |
| HR+ HER2- (Lum. A) | 10 (15,6%) |
| HR+ HER2- (Lum. B) | 11 (17,1%) |
| HR+ HER2+ (Lum. B) | 18 (28,1%) |
| HER2+ HR- | 11 (17,2%) |
| TN | 14 (21,9%) |

Tabelle 15: Aufteilung der molekularen Subgruppen nach visueller und digitaler Analyse des Proliferationsindex (Ki67) (Entdeckungskohorte). Verglichen mit der Auswertung gemäß dem IKWG-Protokoll zeigt die Anwendung von QuPath eine geringere Zuordnung von Fällen zum Luminal B Subtyp.

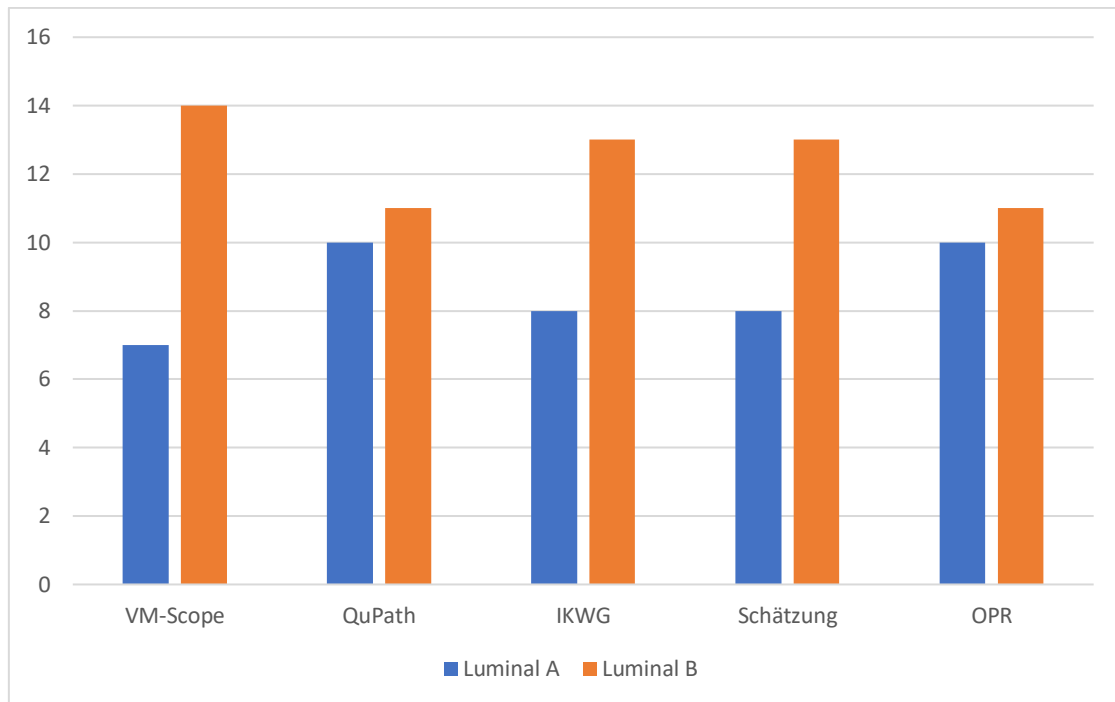


Abbildung 16: Säulendigramm zum Vergleich der unterschiedlichen Auswertungsmethoden des Ki67 Proliferationsindex und den resultierenden Aufteilungen der Fälle zu den molekularen Subtypen Luminal A und B. Im OPR, sowie bei der Verwendung des QuPath-Protokolls werden die Fälle eher dem Luminal A Subtypen zugeordnet. Hingegen klassifizieren alle anderen angewendeten Methoden die Mehrheit der ausgewerteten Fälle in den Subtyp Luminal B.

3.1.3 Visuelle und digitale Auswertung des PhH3 Mitoseindex

Die Untersuchung des Mitoseindex verdeutlicht, dass die Anwendung der PhH3-Färbung im Vergleich zur HE-Färbung eine erheblich höhere Anzahl von erkannten Mitosen ermöglicht. Dieser Befund, wie bereits erwähnt, hat maßgebliche Auswirkungen auf die am Grading orientierte Klassifikation der Gewebeproben. Bei der visuellen Auswertung der PhH3-Färbung (M: 43,8 (Mitosen pro 10 HPF), IQ: 42,6) zeigt sich eine deutlich gesteigerte Mitoserate im Vergleich zur HE-Färbung (M: 9, IQ: 10).

Die digitale Auswertung erfasst sogar eine noch höhere Anzahl von Mitosen auf einer GFF von 1,96 mm² (M: 66,2; IQ: 77,03) als die visuelle Analyse (siehe Abbildung 17).

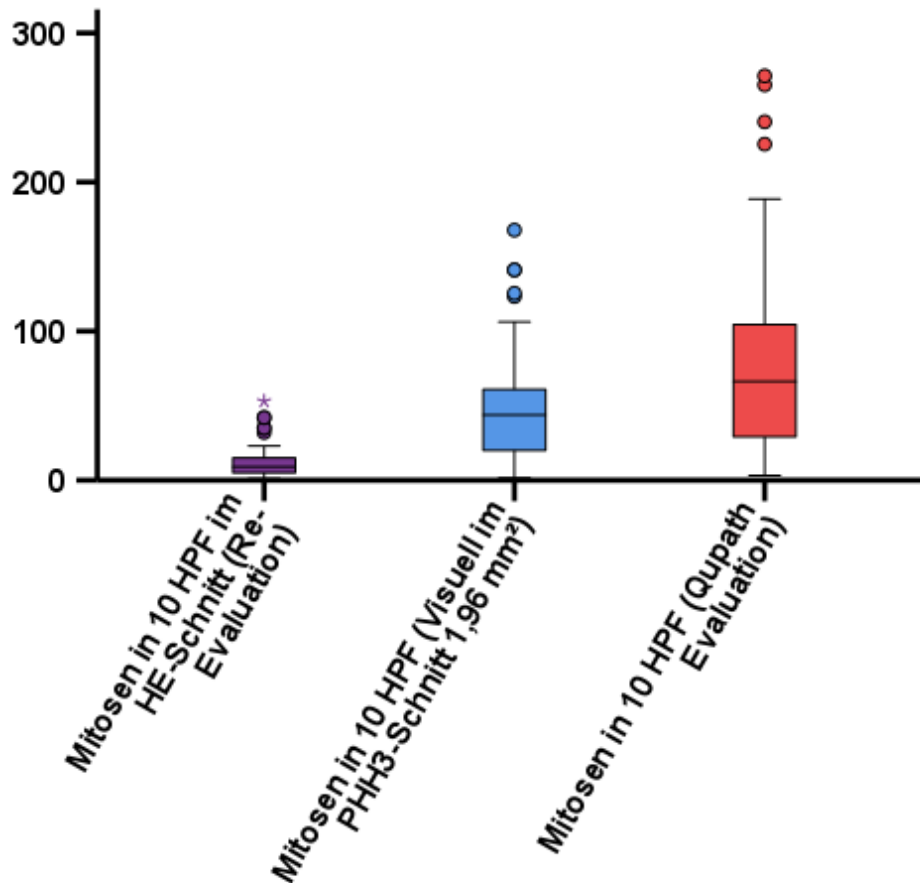


Abbildung 17: Vergleich der Mitoseraten durch visuelle Auswertung von HE- und PhH3 gefärbten Schnitten sowie mithilfe der digitalen Methode (QuPath) (Entdeckungskohorte). Die Anzahl der Mitosen fällt bei digitaler und visueller Auswertung von PhH3 deutlich größer aus als am HE-Schnitt

Bei der Korrelationsanalyse nach Pearson zwischen den einzelnen Methoden zeigt sich ein deutlicher und statistisch signifikanter Zusammenhang ($p < 0,001$). Insbesondere zwischen den digitalen und visuellen PhH3-Methoden manifestiert sich eine ausgeprägte Korrelation ($R: 0,860$).

Die Beziehung zwischen den beiden PhH3-Analysemethoden und der herkömmlichen HE-Auswertung offenbart ebenfalls signifikante Zusammenhänge, wenngleich die Korrelationswerte etwas geringer ausfallen. In diesem Zusammenhang erweist sich die visuelle Methode als minimal stärker korrelierend ($R: 0,675$) verglichen mit der digitalen Methode ($R: 0,635$) (siehe Abbildung 18).

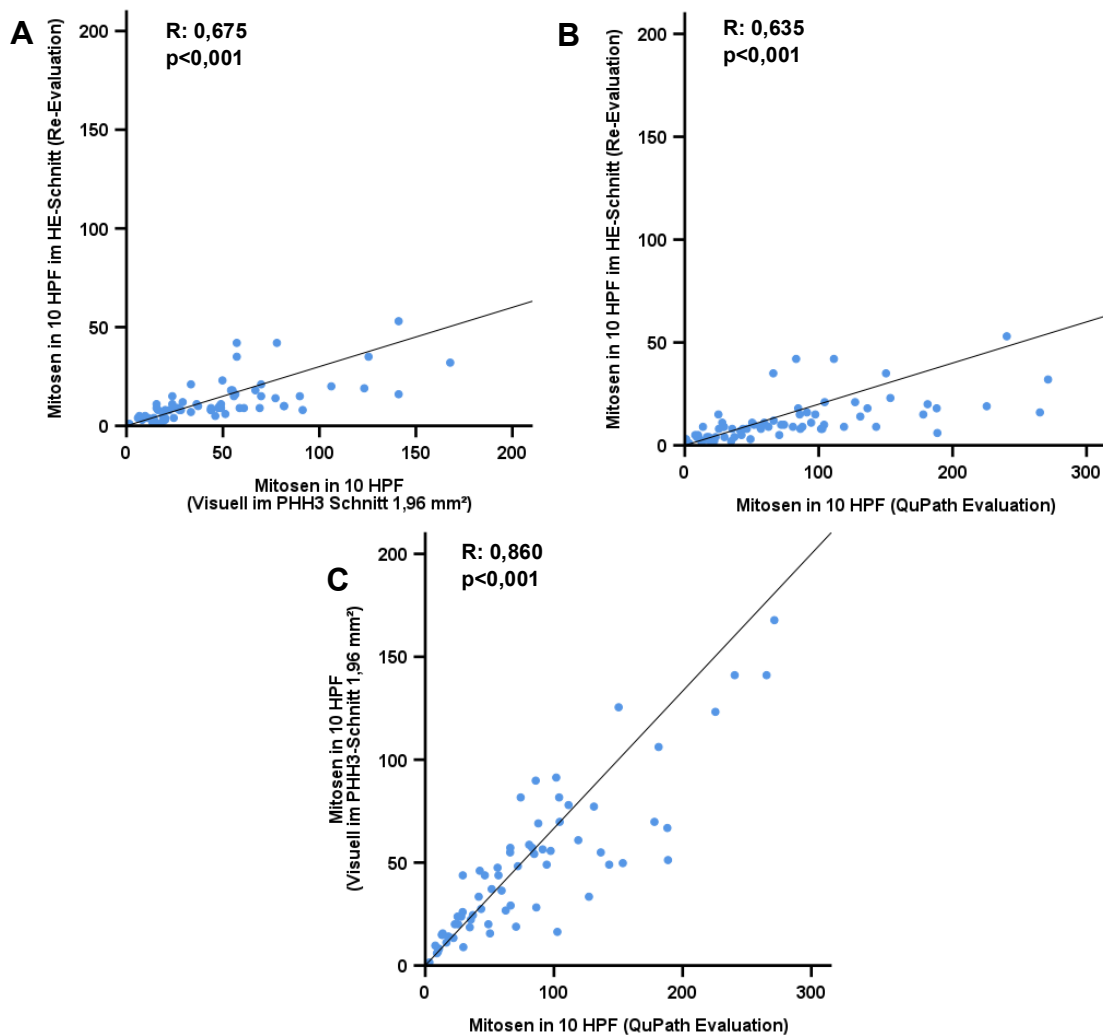


Abbildung 18: Scatter Plots mit den Korrelationen nach Pearson zwischen den einzelnen Mitoseauswertungen (Entdeckungskohorte). (A) HE-Methode und PhH3 visuell, (B) HE-Methode und QuPath und (C) PhH3 visuell und QuPath. Die Korrelation zwischen der HE-Methode und beiden PhH3-Auswertungen fällt geringer aus als zwischen den PhH3-Methoden untereinander.

3.1.4 Clusteranalyse

Im Rahmen des multiparametrischen Ansatzes zur Einstufung des invasiven Mammakarzinoms NST bilden die Parameter tubuläre Formation, nukleärer Pleomorphismus, visuelle Auswertung von PhH3 gefärbten Mitosen und digitale Auswertung des Ki67 Proliferationsindex die Grundlage für die durchgeführte Clusteranalyse. Die Ergebnisse verdeutlichen, dass eine signifikante Neuordnung der meisten Fälle stattgefunden hat.

In diesem Zusammenhang zeigen die Fälle, die ursprünglich mit Grad 1 bewertet wurden, eine Tendenz zur Bildung eines gemeinsamen Clusters. Hingegen verteilen sich die Fälle mit Grad 2 und Grad 3 auf zwei neue Cluster, wie in

Abbildung 19 ersichtlich ist. Diese neu definierten Cluster wurden im Laufe der Untersuchung als „low risk“, „medium risk“ und „high risk“ Cluster bezeichnet.

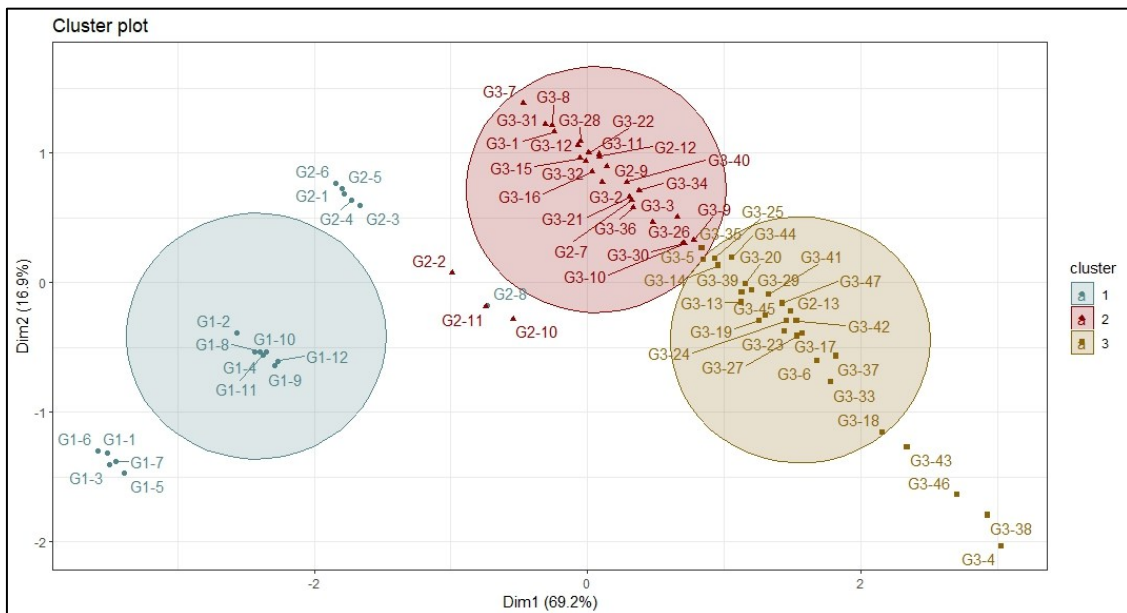


Abbildung 19: Clusteranalyse der Entdeckungskohorte. Das Diagramm zeigt die Neugruppierung der einzelnen Fälle in drei Cluster. (1 blau) low risk, (2 rot) medium risk und (3 gelb) high risk. Insbesondere Fälle mit Grad 2 und 3 werden neu zugeordnet. Die Basis der Clusteranalyse bilden die Parameter: tubuläre Formation, nukleärer Pleomorphismus, visuelle Mitoseauswertung (PhH3) und digitale Auswertung von Ki67.

Innerhalb des „low risk“-Clusters sind hauptsächlich Fälle anzutreffen, die durch niedrige Werte in Bezug auf tubulären Score (M: 2) sowie nukleären Score (M: 2) charakterisiert sind. Zusätzlich weisen sie eine geringe Mitoserate (Mittelwert (MW): 17,4) und einen niedrigeren Proliferationsindex (MW: 13,3%) auf. In den Clustern 2 und 3 sind sowohl die Parameter der tubulären Formation (Cluster 2: M: 3; Cluster 3: M: 3) als auch der Grad des nukleären Pleomorphismus (Cluster 2 und 3: M: 3) fast ausschließlich mit den höchsten Scores vertreten.

Der wesentliche Unterschied zwischen dem zweiten und dem dritten Cluster besteht darin, dass im zweiten Cluster eine geringere Mitoserate festzustellen ist (MW: 46,6) sowie ein reduzierter Proliferationsindex (MW: 32,6%). Demgegenüber fallen die Werte im dritten Cluster für die Mitoserate und den Proliferationsindex höher aus (Mitoserate: MW: 109,6; Proliferationsindex: MW: 69,5%) (siehe Abbildung 20).

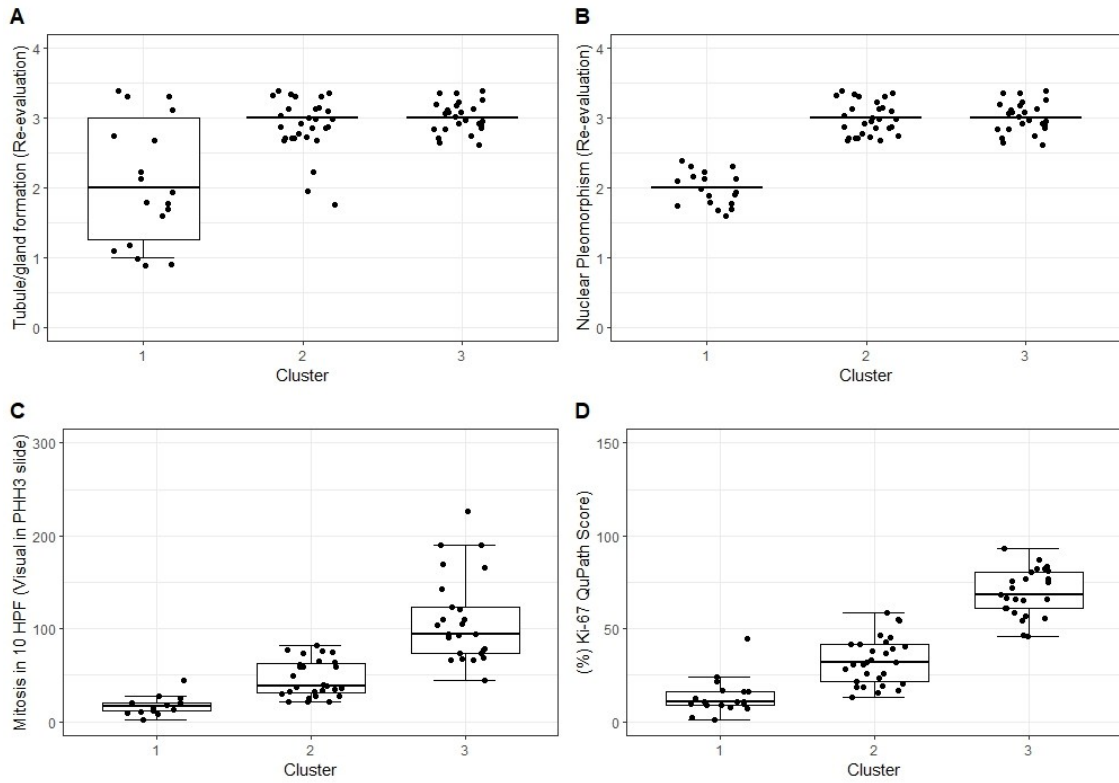


Abbildung 20: Boxplots zur Darstellung der Verteilung der einzelnen Parameter innerhalb der Cluster (Entdeckungkohorte). Die tubulären Formationen (A) und die nukleären Pleomorphismen (B) sind nur in Cluster 1 mit einem niedrigen Score vertreten. Sowohl die PH3-Mitoserate (C) als auch der Ki67-Proliferationsindex (D) sind im ersten Cluster am geringsten und im dritten Cluster am höchsten ausgeprägt.

3.2 Validierungskohorte

3.2.1 Re-Grading

Das Re-Grading der einzelnen Fälle in der Validierungskohorte wurde im Rahmen dieser Studie wie zuvor erwähnt lediglich anhand der Resektatproben durchgeführt. Wie aus Abbildung 21 hervorgeht, führte die erneute Bewertung zu einer erhöhten Einstufung von Fällen zu Grad 3. Dabei wurden 13 Fälle von Grad 2 zu Grad 3 hochgestuft, außerdem ein Fall, der ursprünglich als Grad 1 klassifiziert war. Gleichzeitig erfolgten fünf Abstufungen von Grad 3 zu Grad 2 und eine von Grad 2 zu Grad 1.

Die Auswertung der PhH3-gefärbten Präparate hatte nur eine geringfügige Auswirkung auf die Gesamtanzahl der Fälle pro Grad (siehe Tabelle 16). Dennoch zeigt Abbildung 21, dass einige Fälle neu klassifiziert wurden. In diesem Zusammenhang erfolgte die Herabstufung eines Falles von Grad 2 zu Grad 1 sowie von sieben Fällen von Grad 3 zu Grad 2. Zusätzlich wurden acht Fälle von Grad 2 auf Grad 3 hochgestuft (siehe Abbildung 21).

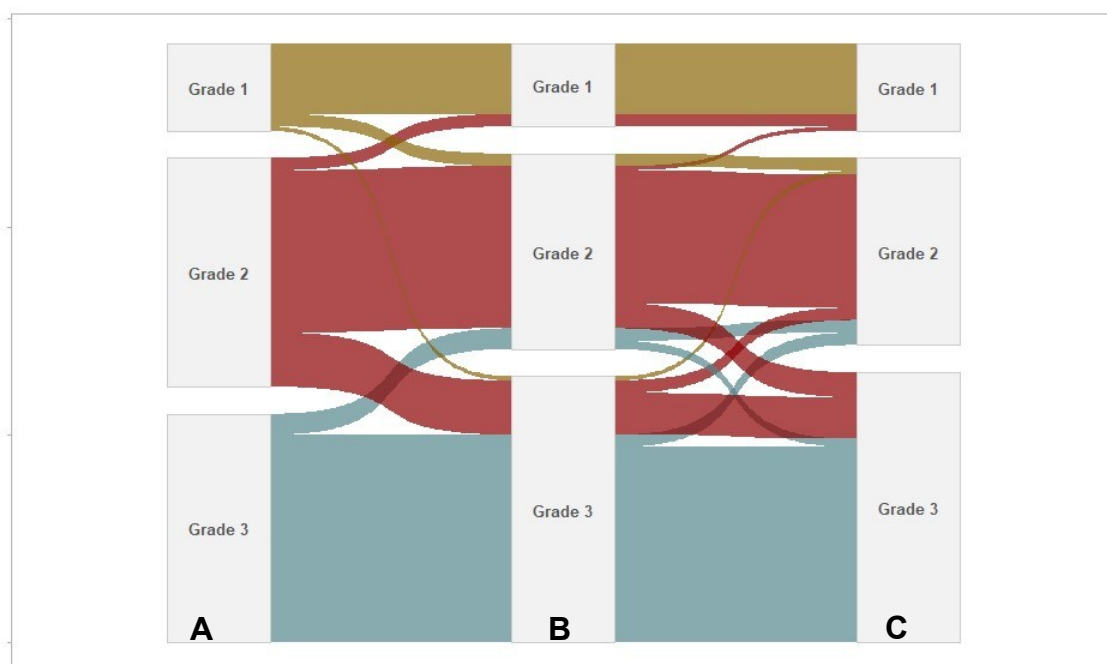


Abbildung 21: Sankey-Plot zur Darstellung der Veränderungen des Gradings nach den einzelnen Auswertungsmethoden (Validierungskohorte). Dabei werden (A) die Original-Gradings, (B) die Re-Evaluation und (C) die Re-Klassifikation nach visueller PhH3-Auswertung aufgeführt. Nach dem Re-Grading und der Re-Klassifikation werden vermehrt Fälle als Grad 3 eingestuft.

| | Gesamt (N=131) |
|--|-----------------------|
| Grading Originalbefund | |
| 1 | 21 (16.0%) |
| 2 | 55 (42.0%) |
| 3 | 55 (42.0%) |
| Re-Evaluation des Gradings | |
| 1 | 20 (15.3%) |
| 2 | 47 (35.9%) |
| 3 | 64 (48.9%) |
| Re-Klassifikation nach visuellem Mitosescore (pHH3) | |
| 1 | 21 (16.0%) |
| 2 | 45 (34.4%) |
| 3 | 65 (49.6%) |

Tabelle 16: Re-Klassifikation nach Re-Evaluation und PhH3-Auswertung (Validierungskohorte). Die Tabelle zeigt die Anpassungen der Klassifikationen nach den einzelnen Re-Evaluationsmethoden. Die Ergebnisse des Re-Gradings anhand von HE-Präparaten und der Re-Klassifikation anhand von PhH3-Präparaten unterscheiden sich nur geringfügig.

3.2.2 Visuelle und digitale Auswertung des Ki67-Proliferationsindex

3.2.2.1 Resektat

Neben den Stanzbiopsien wurden in der Validierungskohorte zusätzlich die Tumorsektate in Hinblick auf ihre Proliferationsrate untersucht. Insgesamt konnten hierbei 129 Proben ausgewertet und verglichen werden. Die Ergebnisse veranschaulichen, dass sowohl die visuelle Schätzung (M: 7,5%, IQ: 18,9%) als auch die Auswertung mittels des IKWG-Protokolls (M: 7,125%, IQ 18,9%) der Resektate ebenfalls sehr ähnliche Resultate erzielen. Darüber hinaus offenbart der Vergleich zur digitalen Methode (M: 5,85%, IQ: 13,1%) lediglich geringfügige Abweichungen (siehe Abbildung 22).

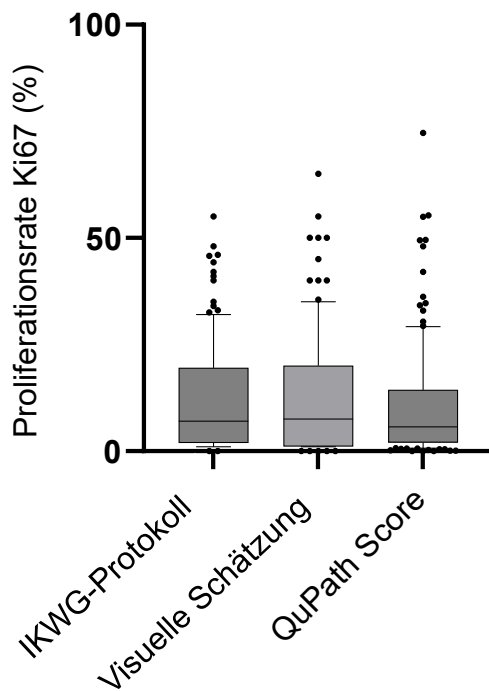


Abbildung 22: Vergleich der Proliferationsindizes zwischen visuellen und digitalen Methoden (Resektate, Validierungskohorte). Die dargestellten Boxplots zeigen eine starke Übereinstimmung der Ergebnisse der visuellen und digitalen Auswertungsmethoden von Ki67.

Bei einer erneuten Untersuchung der einzelnen Proliferationsraten, eingeteilt in Gruppen von 0-5% bis >40%, wird deutlich, dass die Auswertung durch QuPath die geringste Anzahl von Fällen mit einer Proliferationsrate >20% aufweist (QuPath: 11,6%; IKWG: 19,4%; Schätzung: 17,8%). Interessanterweise ist die Diskrepanz zwischen den einzelnen Methoden für Fälle mit >40% positiven Zellen bei den Resektate nicht so ausgeprägt wie bei den Stanzbiopsien der ersten Kohorte (QuPath: 5,4%; IKWG: 5,4%; Schätzung: 4,6%) (siehe Abbildung 23).

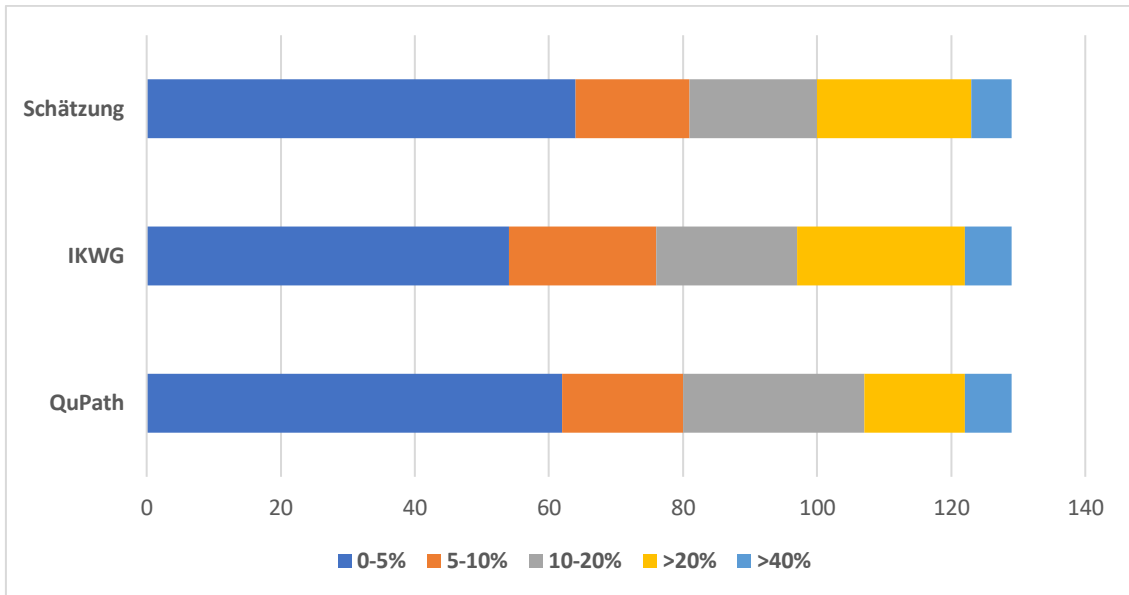


Abbildung 23: Gestapeltes Balkendiagramm mit der Aufteilung der visuellen und digitalen Methoden nach Gruppen unterschiedlicher Proliferationsraten in % (Resektate Validierungskohorte). Die Gruppe der Fälle mit einer Proliferation von >20% ist bei den visuellen Methoden (Schätzung, IKWG-Protokoll) etwas größer als bei der digitalen Auswertung (QuPath).

Auch in der Validierungskohorte zeigen sich statistisch signifikante Korrelationen (Pearson) zwischen den einzelnen angewandten Methoden ($p < 0,001$) (siehe Abbildung 24). Erneut manifestiert sich die stärkste Korrelation zwischen den beiden visuellen Methoden ($R: 0,983$).

Im Vergleich zwischen der digitalen und der visuellen Methode zeigt sich, dass die visuelle Schätzung minimal stärker mit der Bewertung durch QuPath korreliert ($R: 0,879$) als die Anwendung des IKWG-Protokolls ($R: 0,862$).

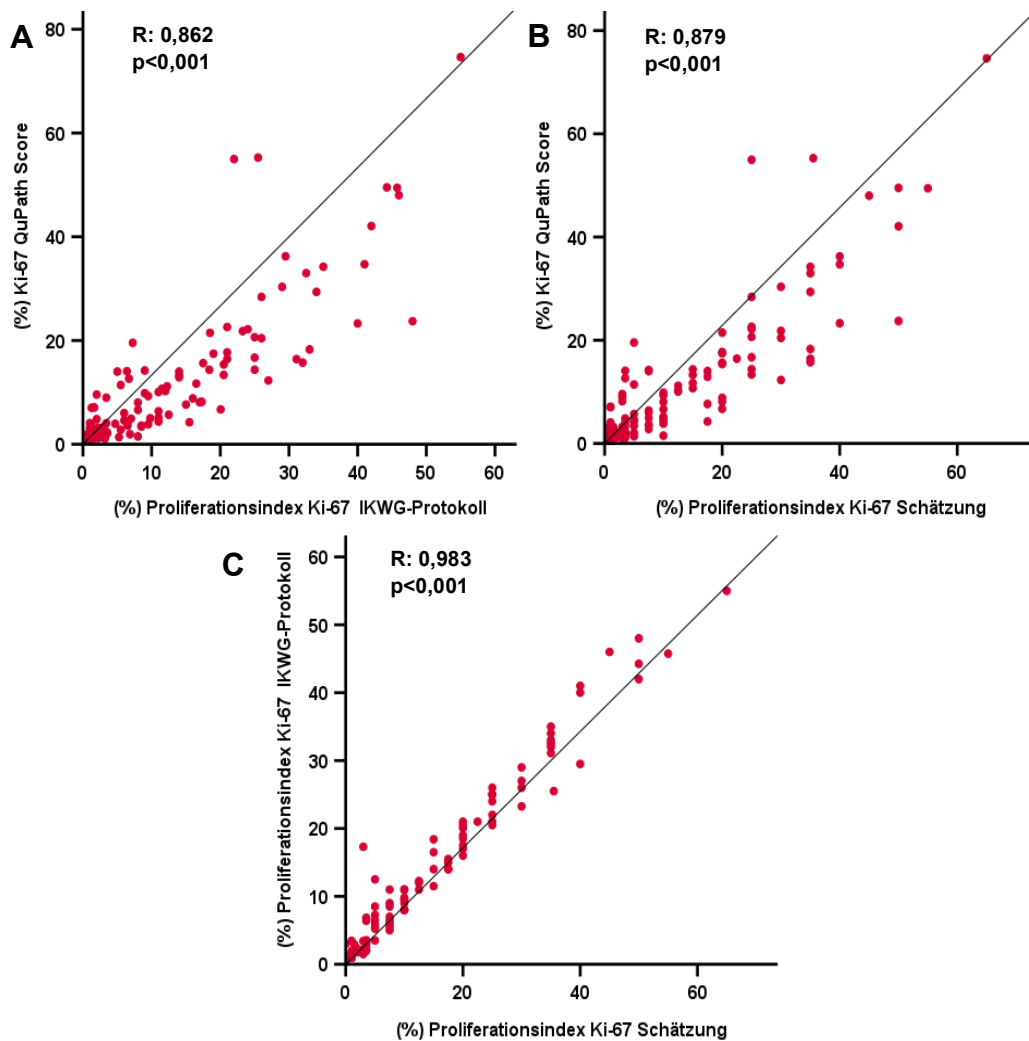


Abbildung 24: Scatter-Plots zur Darstellung der Korrelation nach Pearson der verwendeten Auswertungsmethoden des Ki67-Proliferationsindex (Resektate, Validierungskohorte). Alle Methoden korrelieren hierbei signifikant miteinander. (A) Die Auswertung mit dem IKWG-Protokoll korreliert etwas geringer mit der digitalen Methode (QuPath) als die visuelle Schätzung (B). Die stärkste Korrelation zeigt sich zwischen den beiden visuellen Methoden (C).

3.2.2.1.1 Veränderung des molekularen Subtyps nach Auswertung der Resektate

Die Analyse der Resektatproben aus der Entdeckungskohorte zeigt in Bezug auf die Re-Klassifikation des molekularen Gradings, dass sowohl visuelle als auch digitale Verfahren die Mehrheit der Fälle als Luminal A einstufen (Schätzung: 42,6%; IKWG-Protokoll: 44,2%; QuPath: 46,5%). Weiterhin verdeutlichen die Ergebnisse, ähnlich wie in der Entdeckungskohorte, dass die digitale Methode im Vergleich zu den visuellen Methoden zu einer geringeren Klassifikation von Fällen als Luminal B führt (QuPath: 17,8%; Schätzung: 21,7%; IKWG-Protokoll: 20,1%) (siehe Tabelle 17).

| | Gesamt (N=129) |
|---|---------------------------|
| Molekularer Subtyp nach Schätzung | |
| HR+ HER2- (Lum. A) | 55 (42,6%) |
| HR+ HER2- (Lum. B) | 28 (21,7%) |
| HR+ HER2+ (Lum. B) | 16 (12,4%) |
| HER2+ HR- | 5 (3,9%) |
| TN | 25 (19,4%) |
| Molekularer Subtyp nach IKWG-Protokoll | |
| HR+ HER2- (Lum. A) | 57 (44,2%) |
| HR+ HER2- (Lum. B) | 26 (20,1%) |
| HR+ HER2+ (Lum. B) | 16 (12,4%) |
| HER2+ HR- | 5 (3,9%) |
| TN | 25 (19,4%) |
| Molekularer Subtyp nach QuPath | |
| HR+ HER2- (Lum. A) | 60 (46,5%) |
| HR+ HER2- (Lum. B) | 23 (17,8%) |
| HR+ HER2+ (Lum. B) | 16 (12,4%) |
| HER2+ HR- | 5 (3,9%) |
| TN | 25 (19,4%) |

Tabelle 17: Aufteilung der molekularen Subgruppen nach visueller und digitaler Analyse des Proliferationsindex (Ki67) am Resektat (Validierungskohorte). Verglichen mit der Auswertung gemäß dem IKWG-Protokoll und der Schätzung des Proliferationsindex zeigt die Anwendung von QuPath eine geringere Zuordnung von Fällen zum Luminal B Subtyp.

3.2.2.2 Stanzbiopsie

Die Analyse der Stanzbiopsien zeigt ähnliche Ergebnisse wie die Auswertung der Resektate. Hierbei konnten 125 Proben ausgewertet und verglichen werden. Sowohl die visuelle Schätzung (M: 16,25%, IQ: 22,5%) und die Auswertung nach IKWG-Protokoll (M: 15,75%, IQ: 20,5%) weisen lediglich minimale Unterschiede in ihren Ergebnissen auf. Die Auswertung mittels QuPath (M: 13,57%, IQ: 22,01%) erbringt einen leicht geringeren Median als die visuellen Methoden (siehe Abbildung 27).

Teilt man die Fälle erneut nach Gruppen von 0-5% bis >40% ein so wird deutlich, dass mittels QuPath zwar deutlich weniger Fälle in die Kategorie >20% eingestuft werden (QuPath: 20,8%; IKWG: 29,6%; Schätzung: 28,8%), jedoch durch die digitale Auswertung eindeutig mehr Fälle mit einer Proliferationsrate von >40% identifiziert werden (QuPath: 16,8%; IKWG: 9,6%; Schätzung: 10,4%) (siehe Abbildung 25).

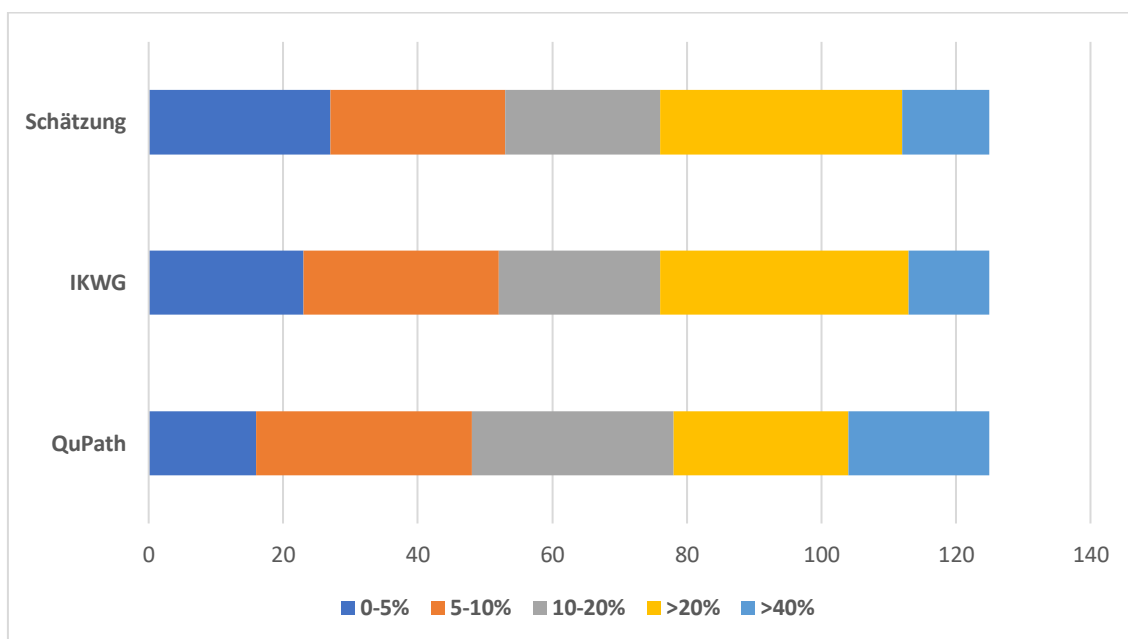


Abbildung 25: Gestapeltes Balkendiagramm mit der Aufteilung der visuellen und digitalen Methoden nach Gruppen unterschiedlicher Proliferationsraten in % (Resektate Validierungskohorte). Die Gruppe der Fälle mit einer Proliferationsrate von >20% ist bei den visuellen Methoden (Schätzung, IKWG-Protokoll) etwas größer als bei der digitalen Auswertung (QuPath). Jedoch ordnet die digitale Auswertung deutlich mehr Fällen eine Ki67-Proliferationsrate von >40% zu (21 Fälle QuPath, 12 Fälle IKWG).

Die wiederholte Gegenüberstellung aller angewandten Methoden ergibt signifikante Korrelationen untereinander ($p < 0,001$). Dabei zeigt sich erneut die stärkste Korrelation zwischen den visuellen Methoden ($R: 0,993$). Die Bewertung nach IKWG-Protokoll korreliert etwas besser mit der digitalen Methode ($R: 0,906$) im Vergleich zur visuellen Schätzung ($R: 0,899$) (siehe Abbildung 26).

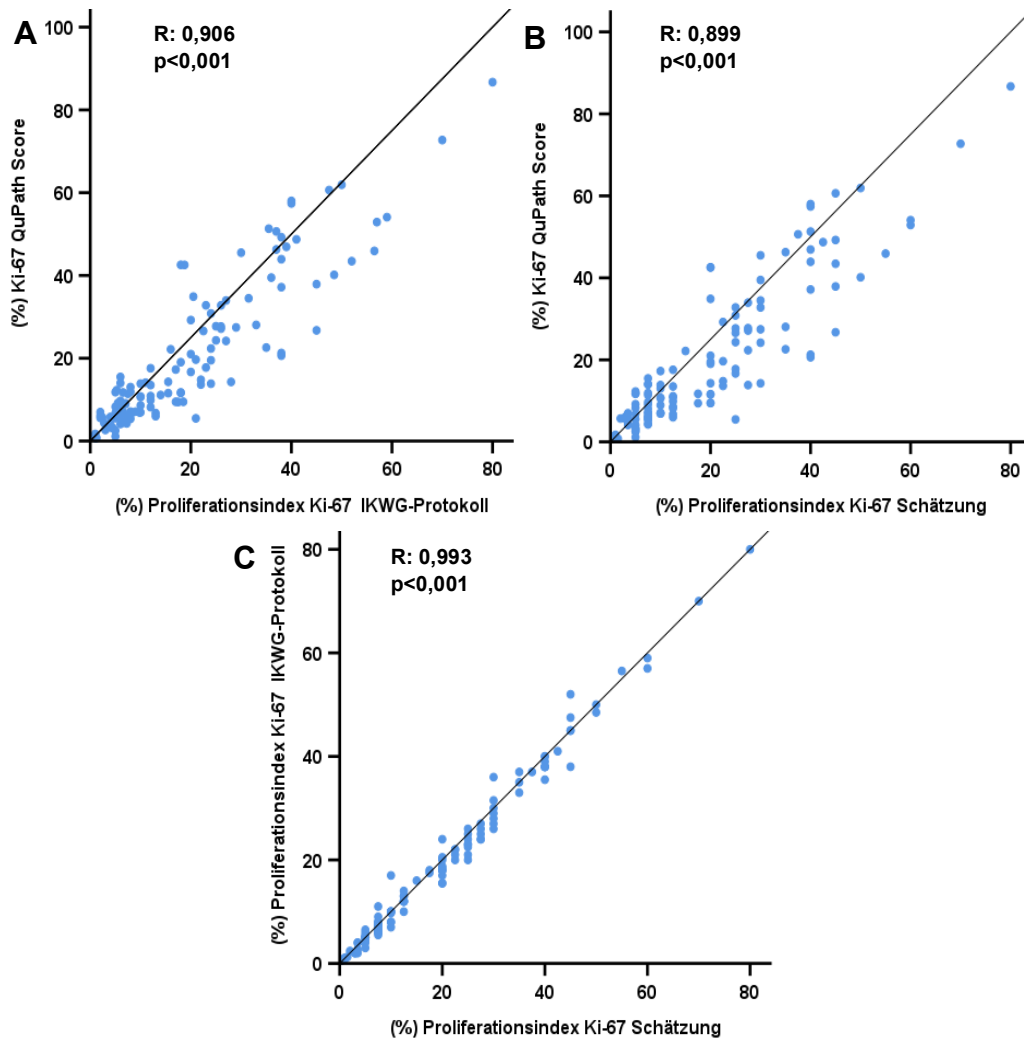


Abbildung 26: Scatter-Plots zur Darstellung der Korrelation nach Pearson der verwendeten Auswertungsmethoden des Ki67-Proliferationsindex (Stanzbiopsien, Validierungskohorte). Alle Methoden korrelieren hierbei signifikant miteinander. (A) Die Auswertung mit dem IKWG-Protokoll korreliert etwas stärker mit der digitalen Methode (QuPath) als die visuelle Schätzung (B). Die stärkste Korrelation zeigt sich zwischen den beiden visuellen Methoden.

3.2.2.2.1 Veränderung des molekularen Subtyps nach Auswertung der Stanzbiopsien

Die Analyse der Stanzbiopsien zeigt eine vergleichbare Verteilung von Luminal A und Luminal B Fällen bei der visuellen Schätzung und der Auswertung des Ki67-Proliferationsindex mittels des IKWG-Protokolls (Schätzung: Luminal A: 32%, Luminal B: 34,4%; IKWG-Protokoll: Luminal A: 32,8%, Luminal B: 33,6%). Im Gegensatz dazu führt die digitale Auswertung zu einer deutlich geringeren Anzahl an Luminal B Fällen (26,4%). Diese Ergebnisse stimmen mit der in der Entdeckungskohorte sowie in den Resektaten festgestellten Tendenz überein, dass die digitale Auswertung dazu neigt, mehr Fälle als Luminal A zu klassifizieren (siehe Tabelle 18).

| | Gesamt (N=125) |
|---|---------------------------|
| Molekularer Subtyp nach Schätzung | |
| HR+ HER2- (Lum. A) | 40 (32%) |
| HR+ HER2- (Lum. B) | 43 (34,4%) |
| HR+ HER2+ (Lum. B) | 14 (11,2%) |
| HER2+ HR- | 3 (2,4%) |
| TN | 25 (20%) |
| Molekularer Subtyp nach IKWG-Protokoll | |
| HR+ HER2- (Lum. A) | 41 (32,8%) |
| HR+ HER2- (Lum. B) | 42 (33,6%) |
| HR+ HER2+ (Lum. B) | 14 (11,2%) |
| HER2+ HR- | 3 (2,4%) |
| TN | 25 (20%) |
| Molekularer Subtyp nach QuPath | |
| HR+ HER2- (Lum. A) | 50 (40%) |
| HR+ HER2- (Lum. B) | 33 (26,4%) |
| HR+ HER2+ (Lum. B) | 14 (11,2%) |
| HER2+ HR- | 3 (2,4%) |
| TN | 25 (20%) |

Tabelle 18: Aufteilung der molekularen Subgruppen nach visueller und digitaler Analyse des Proliferationsindex (Ki67) an den Stanzbiopsien (Validierungskohorte). Verglichen mit der Auswertung gemäß dem IKWG-Protokoll und der Schätzung des Proliferationsindex zeigt die Anwendung von QuPath eine geringere Zuordnung von Fällen zum Luminal B Subtyp.

3.2.2.3 Stanzbiopsie vs. Resektat

Bei der Untersuchung der verschiedenen Probenarten wird in dieser Studie deutlich, dass Stanzbiopsien im Durchschnitt höhere Proliferationsraten aufweisen als Resektate (Schätzung: 20,05% vs. 12,79%, IKWG-Protokoll: 19,43% vs. 12,13%, QuPath: 20,23% vs. 11,12%) (siehe Abbildung 27).

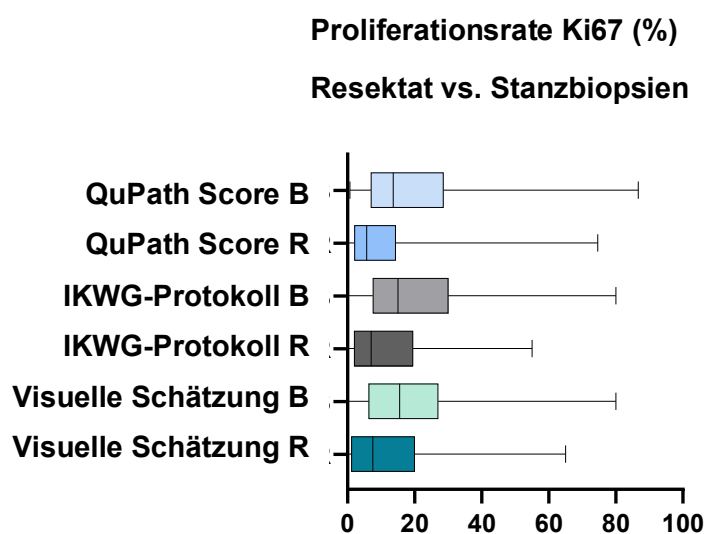


Abbildung 27: Vergleich der Ki67-Proliferationsindizes zwischen den einzelnen Auswertungsmethoden an Stanzbiopsien und Resektaten (Validierungskohorte). Die Boxplots zeigen, dass alle Methoden für Stanzbiopsien deutlich höhere Proliferationsraten ergeben als für Resektate.

Gleichzeitig offenbaren die Ergebnisse der einzelnen Methoden zwar signifikante ($p: 0,0001$) jedoch lediglich moderate Korrelationen. Die engsten Korrelationen bestehen zwischen den Ergebnissen der Stanzbiopsien und denen der Resektate bei der Auswertung mittels QuPath ($R: 0,41$) sowie durch das IKWG-Protokoll ($R: 0,40$). Die geringste Korrelation wird bei der visuellen Schätzung der Proliferationsrate festgestellt ($R: 0,33$) (siehe Abbildung 28).

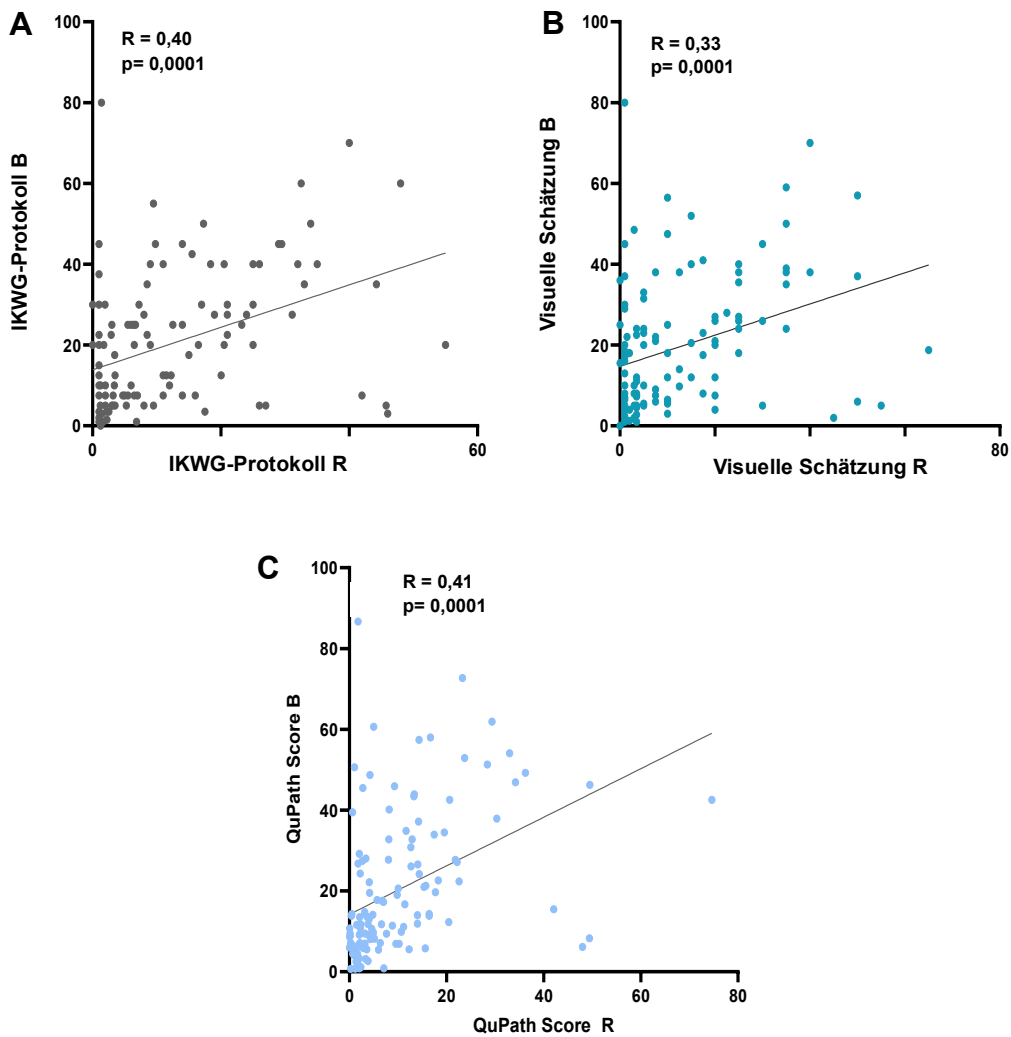


Abbildung 28: Scatter-Plots zur Darstellung der Korrelation nach Pearson der verwendeten Auswertungsmethoden des Ki67 Proliferationsindex zwischen Resektaten und Stanzbiopsien. Es zeigt sich eine signifikante, jedoch nur moderate Korrelation zwischen den Proliferationsraten von Stanzbiopsien und Resektaten. Die einzelnen Methoden: (A) IKWG-Protokoll, (B) visuelle Schätzung und (C) QuPath unterscheiden sich dabei nur geringfügig.

3.2.2.3.1 Vergleich der molekularen Klassifikation von Stanzbiopsien und Resektaten anhand von visuellen und digitalen Methoden

Beim Vergleich der Klassifikationen der molekularen Subtypen zwischen Stanzbiopsien und Resektaten fällt auf, dass letztere in dieser Studie häufiger dem Luminal A Subtyp zugeordnet werden. Sowohl in der digitalen (Resektate: Luminal A: 46,5%, Luminal B: 17,8% ; Stanzbiopsien: Luminal A: 40%, Luminal B: 26,4%) als auch in den visuellen Auswertungen mittels Schätzung (Resektate: Luminal A: 42,6%, Luminal B: 21,7%; Stanzbiopsien: Luminal A: 32%, Luminal B:

34,4%) sowie gemäß dem IKWG-Protokoll (Resektate: Luminal A: 44,2%, Luminal B: 20,1%; Stanzbiopsien: Luminal A: 32,8%, Luminal B: 33,6%) zeigt sich dieser Trend. Darüber hinaus verdeutlichen die Ergebnisse, dass die digitale Auswertung sowohl bei Stanzbiopsien als auch bei Resektaten tendenziell eher zu einer Klassifikation als Luminal A neigt, im Gegensatz zur visuellen Auswertung (siehe Tabelle 17/18, Abbildung 29).

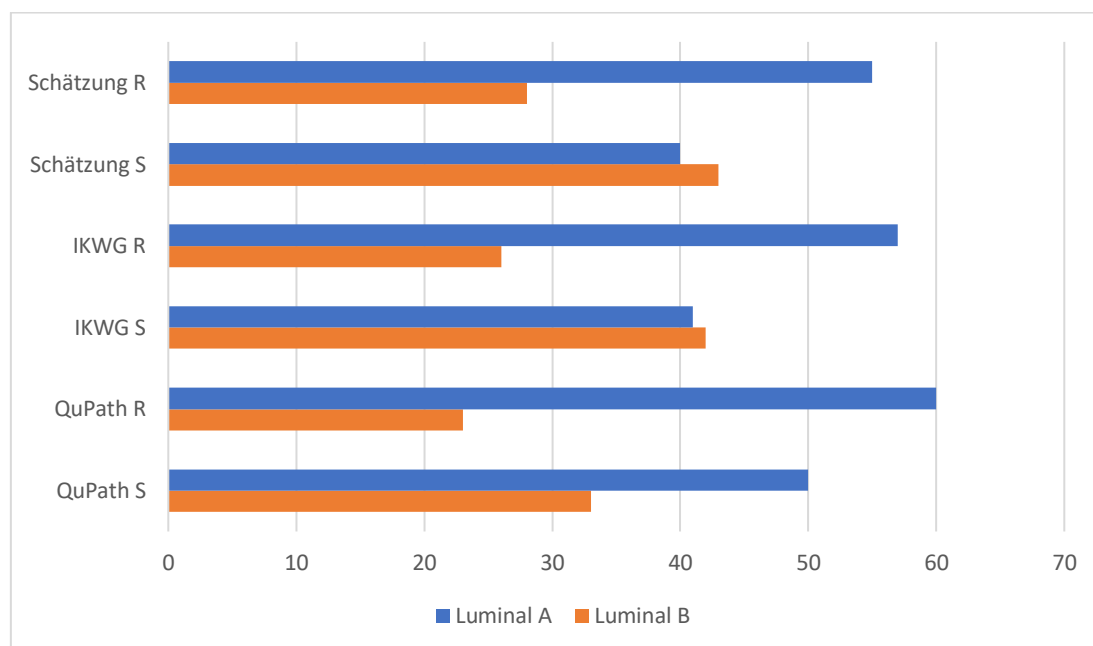


Abbildung 29: Balkendiagramm zur Veranschaulichung der Aufteilung der molekularen Subtypen bei Resektaten und Stanzbiopsien unter Anwendung visueller und digitaler Auswertungsmethoden. Auffallend ist die verstärkte Zuordnung zum Luminal A Subtyp bei der Auswertung von Resektaten durch (A) visuelle Schätzung und (B) mithilfe des IKWG-Protokolls. Zudem erfolgt bei beiden Probenarten vermehrt eine Zuordnung zu diesem Subtyp, wenn die (C) digitale Auswertungsmethode (QuPath) angewendet wird.

3.2.3 Visuelle und digitale Auswertung des PhH3 Mitoseindex

3.2.3.1 Visuelle Auswertung PhH3 vs. HE

Da in dieser Studie das Re-Grading des HE-Schnittes zu diesem Zeitpunkt ausschließlich an den Resektaten durchgeführt wurde, da hier die Tumorerogenität besser erfasst wurde, können die Mitoseraten nur zwischen den verschiedenen histologischen Färbungen (HE, PhH3) dieser Gewebeproben verglichen werden. Wie schon in Kapitel 3.2.1 angedeutet, fällt das Grading nach

PhH3-Auswertung am Resektat ähnlich wie das der Re-Evaluation der HE-Schnitte aus (siehe Abbildung 21, Tabelle 16).

Bei genauerer Betrachtung der Resultate der Mitosezählung zeigt sich, dass sich die Auswertung von HE- (M: 8, IQ: 15) und PhH3-Präparaten (M: 10, IQ: 15) lediglich geringfügig unterscheidet (siehe Abbildung 30).

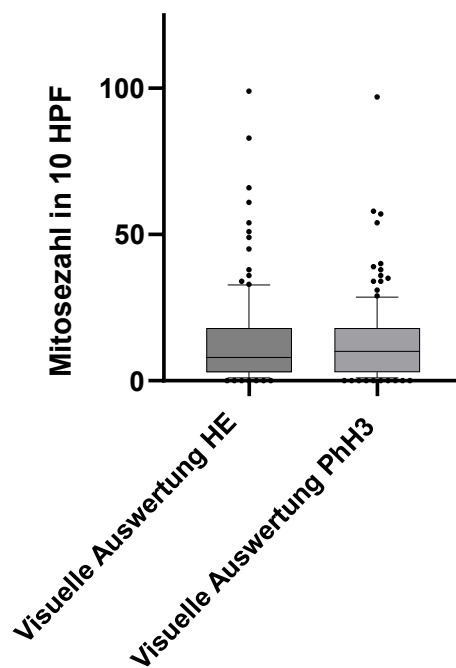


Abbildung 30: Vergleich der Mitoseraten anhand von HE- und PhH3-Färbung. Die Boxplots zeigen für die Mitoseauswertung von Resektaten ähnliche Werte für HE- und PhH3-gefärbte Proben.

Für die beiden Methoden konnte in dieser Studie eine starke signifikante Korrelation festgestellt werden ($p: 0,0001$, $R: 0,78$) (siehe Abbildung 31)

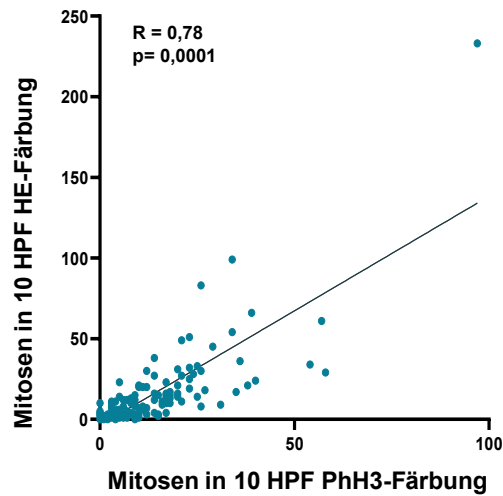


Abbildung 31: Scatter-Plot zur Darstellung der Mitoseraten anhand von HE- und PhH3-gefärbten Schnitten (Resektat). Die Auswertung von Mitosen in 10 HPF zeigt für Resektate bei der Auswertung von HE- und PhH3 gefärbten Schnitten eine starke Korrelation nach Pearson auf.

3.2.3.2 Digitale Auswertung vs. visuelle Auswertung des PhH3

Mitoseindex

Eine digitale Auswertung des MAI am PhH3-Präparat konnte in dieser Studie lediglich an 45 Resektaten und 48 Stanzbiopsien mit dem Grading 1 und 2 durchgeführt werden. Aus den verfügbaren Daten wird ersichtlich, dass die digitale Methode tendenziell höhere Mitosezahlen ermittelt. Dabei zeigt sich, dass der Unterschied zwischen erfassten Mitosezahlen beim Resektat (Visuell: M: 3, IQ: 5; digital: M: 3,36, IQ: 7,69) geringer ausfällt als bei der Auswertung der Stanzbiopsien (Visuell: M: 3, IQ: 5; digital: M: 9,4, IQ: 19,58) (siehe Abbildung 32).

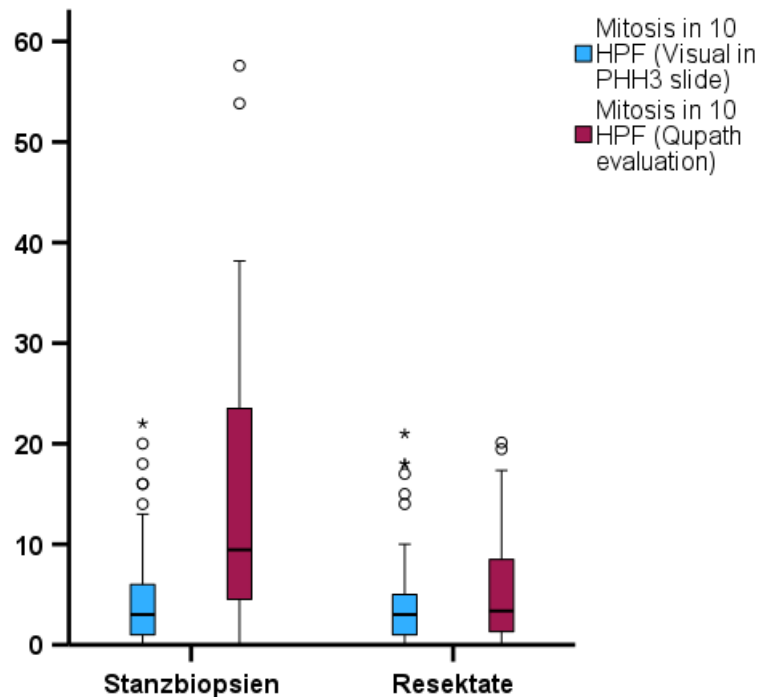


Abbildung 32: Vergleich von digitaler und visueller Auswertung der Mitoseraten von Resektaten und Stanzbiopsien (PhH3). Die Boxplots verdeutlichen, dass die digitale Auswertung in dieser Studie höhere Mitoseraten in 10 HPF ergibt als die visuelle Auswertung. Besonders bei den Stanzbiopsien ist der Unterschied am deutlichsten.

Die Korrelationsanalyse nach Pearson der digitalen und visuellen Methode für beide Probenarten ergibt ein statistisch signifikantes Ergebnis ($<0,001$). Dennoch zeigt sich eine stärkere Korrelation zwischen den beiden Auswertungsarten im Fall von Stanzbiopsien ($R: 0,806$), im Vergleich zu den Resektaten ($R: 0,608$) (siehe Abbildung 33).

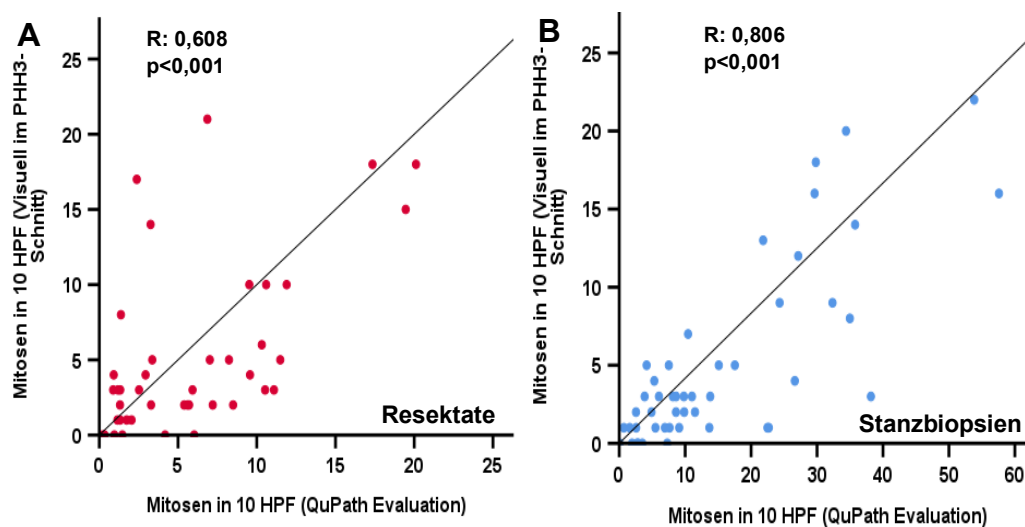


Abbildung 33: Scatterplots zur Darstellung der Korrelation nach Pearson von digitaler und visueller PhH3-Auswertung bei Resektaten und Stanzbiopsien. Die Korrelation der beiden Methoden ist für Resektate (A) geringer als für Stanzbiopsien (B).

3.2.3.3 Stanzbiopsie vs. Resektat

Die Mitoserate der Stanzbiopsien anhand von PhH3 gefärbten Schnitten (M: 12,5, IQ: 22) fällt im Gegensatz zu den Resektaten (M: 10, IQ: 15) höher aus (siehe Abbildung 34).

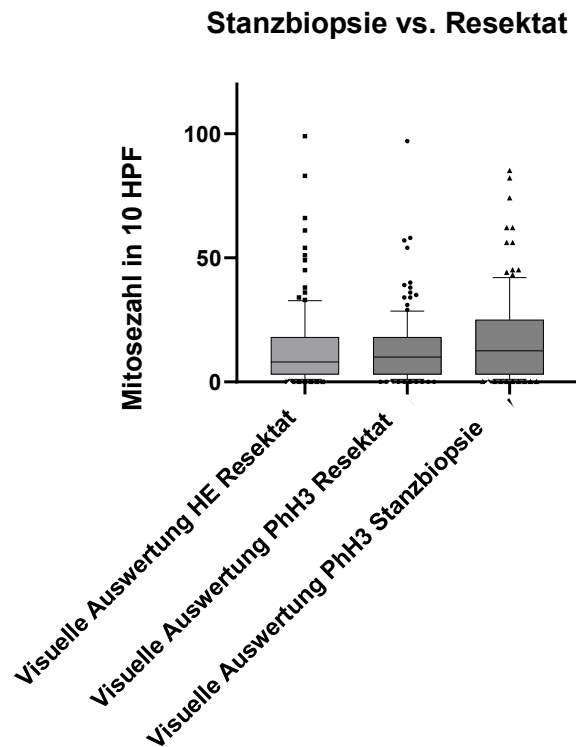


Abbildung 34: Vergleich der visuellen Auswertung der Mitoseraten von Resektaten und Stanzbiopsien (HE, PhH3). Die Mitosezahl in 10 HPF fällt bei PhH3-gefärbten Stanzbiopsien höher aus als bei Resektaten (HE, PhH3)

Zwischen der Auswertung von PhH3 gefärbten Resektaten und Stanzbiopsien zeigt sich ein signifikanter Zusammenhang ($p: 0,0001$, $R: 0,76$). Hinsichtlich der HE-gefärbten Resektate besteht für die PhH3-gefärbten Stanzbiopsien eine moderate und signifikante Korrelation nach Pearson ($p: 0,0001$, $R: 0,37$) (siehe Abbildung 35).

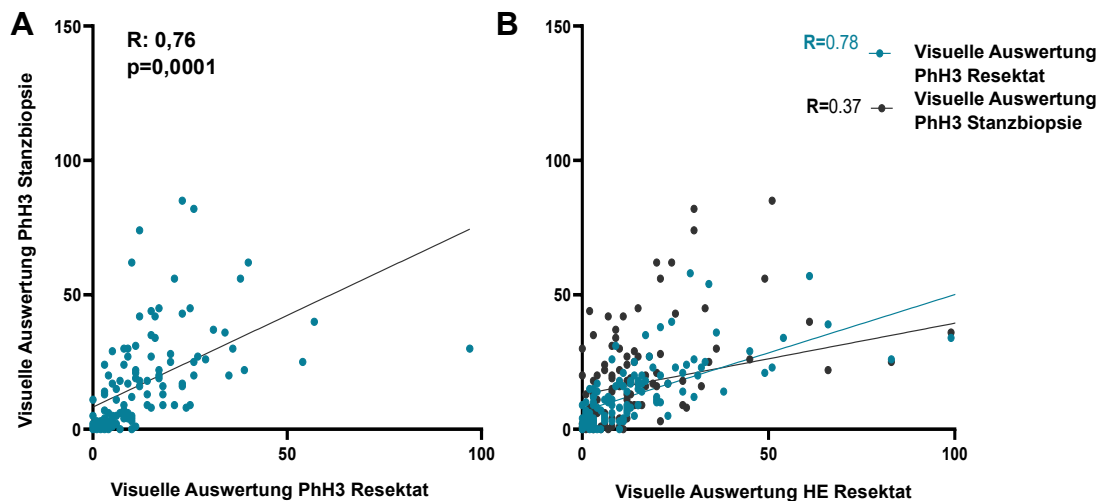


Abbildung 35: Scatterplots zur Darstellung der Korrelation nach Pearson der visuellen Auswertungsmethoden für die Mitoseraten von Stanzbiopsien und Resektaten (HE, PhH3). Die Korrelation zwischen PhH3-gefärbten Stanzbiopsien und Resektaten (A) sowie zwischen PhH3 und HE-gefärbtem Resektat ist größer als zwischen PhH3-gefärbten Stanzbiopsien und HE-gefärbten Resektaten (B).

3.2.4 Clusteranalyse

In dieser Kohorte wurden erneut die vier Parameter tubuläre Formation, nukleärer Pleomorphismus, visuell ausgewertete Mitoserate anhand der PhH3-Färbung und der digital aufgewertete Ki67-Proliferationsindex, betrachtet. Es wurden jedoch Werte der Resektate (tubuläre Formationen, nukleärer Pleomorphismus) und Stanzbiopsien (PhH3 visuell, Ki67 digital) zusammengefasst, da die Aussagekraft der Immunhistochemie bei den Stanzbiopsien höher zu sein scheint. Hierbei ergibt sich, ähnlich wie bei der Entdeckungskohorte, einer Clusterbildung in drei Gruppen. Deutlich wird erneut, dass vor allem Fälle mit Grad 2 und Grad 3 sich neu gruppieren, während die Fälle mit Grad 1 größtenteils ein eigenes Cluster bilden (siehe Abbildung 36).

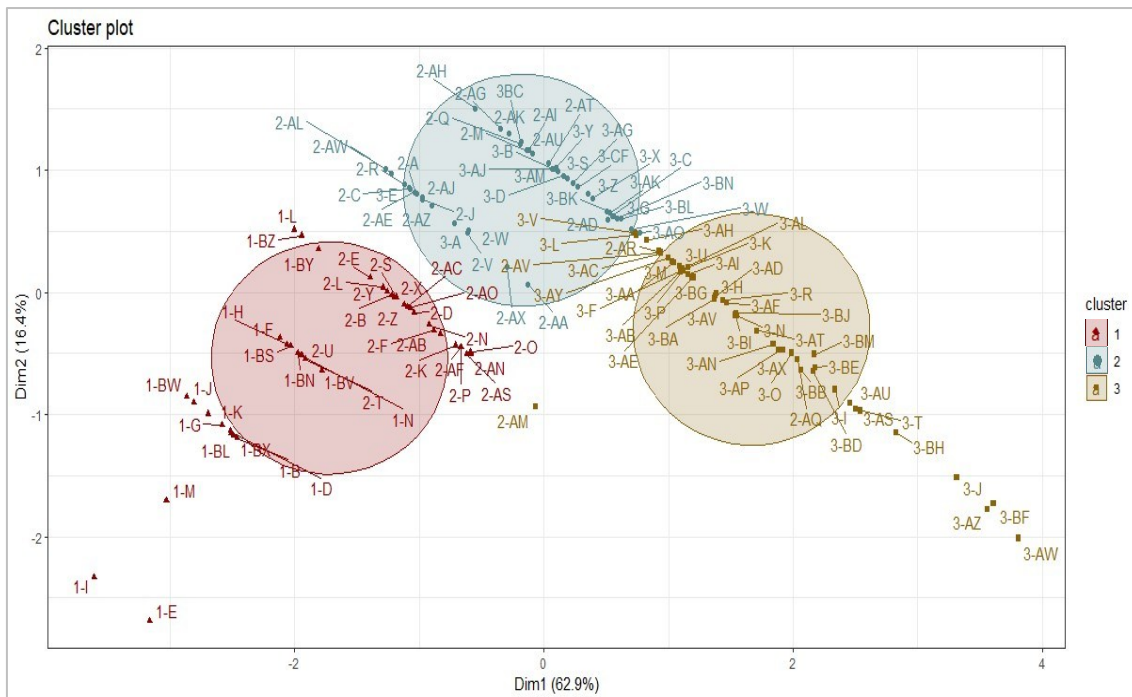


Abbildung 36: Clusteranalyse der Validierungskohorte. Das Diagramm zeigt die Neugruppierung der einzelnen Fälle in drei Cluster. (1 rot) low risk, (2 blau) medium risk, (3 gelb) high risk. Besonders Fälle des Grades 2 und 3 teilen sich neu auf. Die Basis der Clusteranalyse bilden die Parameter: tubuläre Formation (Resektat), nukleärer Pleomorphismus (Resektat), visuelle Mitoseauswertung (PhH3) (Stanzbiopsie) und digitale Auswertung von Ki67 (Stanzbiopsie).

Die tubulären Formationen zeigen im ersten Cluster niedrige Werte (M: 2), ebenso die nukleären Pleomorphismen (M: 2). Im Gegensatz dazu weisen sowohl das zweite als auch das dritte Cluster größtenteils die maximalen Scores für diese beiden Parameter auf (M: 3). Die visuell am PhH3-Präparat ermittelte Mitoserate pro 10 HPF ist im „low risk“-Cluster am niedrigsten (MW: 3,8), gefolgt vom „medium“- (MW: 10,6) und dem „high risk“-Cluster (MW: 37,1). Eine ähnliche Tendenz zeigt sich bei den digital ermittelten Proliferationsraten (QuPath) der einzelnen Gruppen. Auch hier weist das erste Cluster im Durchschnitt die niedrigste Proliferationsrate auf (MW: 8,2%, QuPath digital), gefolgt vom zweiten Cluster (MW: 12,3%), während das dritte Cluster die mit Abstand höchste durchschnittliche Proliferationsrate zeigt (MW: 40,7%) (siehe Abbildung 37).

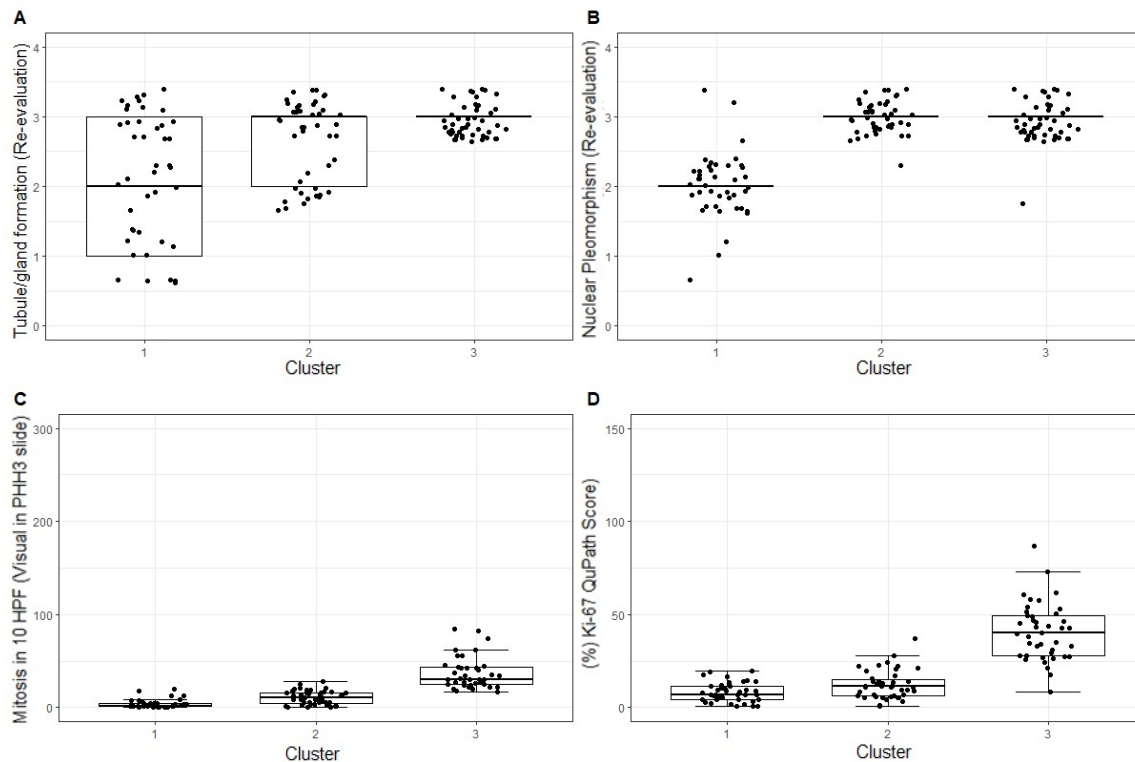


Abbildung 37: Boxplots zur Darstellung der Verteilung der einzelnen Parameter innerhalb der Cluster. Die tubulären Formationen (A) und die nukleären Pleomorphismen (B) sind nur in Cluster 1 mit einem niedrigen Score vertreten. Sowohl PhH3-Mitoseraten (C) als auch der Ki67-Proliferationsindex (D) sind im ersten Cluster am geringsten und im dritten Cluster am höchsten ausgeprägt.

3.2.5 Korrelation der Clusteranalyse zum klinischen Follow-Up

3.2.5.1 Korrelation mit dem rezidivfreien Überleben (RFS)

Im Rahmen dieser Studie konnte kein signifikanter Zusammenhang zwischen dem Grading ($p: 0,4917$) oder dem multiparametrischen Cluster ($p: 0,5406$) in Bezug auf das rezidivfreie Überleben hergestellt werden. Beim konventionellen Grading ist dennoch in den Kurven ein Trend für besseres Überleben von G1-Karzinomen erkennbar. Darüber hinaus zeigen die entsprechenden Kaplan-Meyer-Diagramme aber keine deutliche Abgrenzung voneinander, was darauf hinweist, dass die Raten des rezidivfreien Überlebens zwischen den verschiedenen Graden und Clustern nicht wesentlich variieren (siehe Abbildung 38).

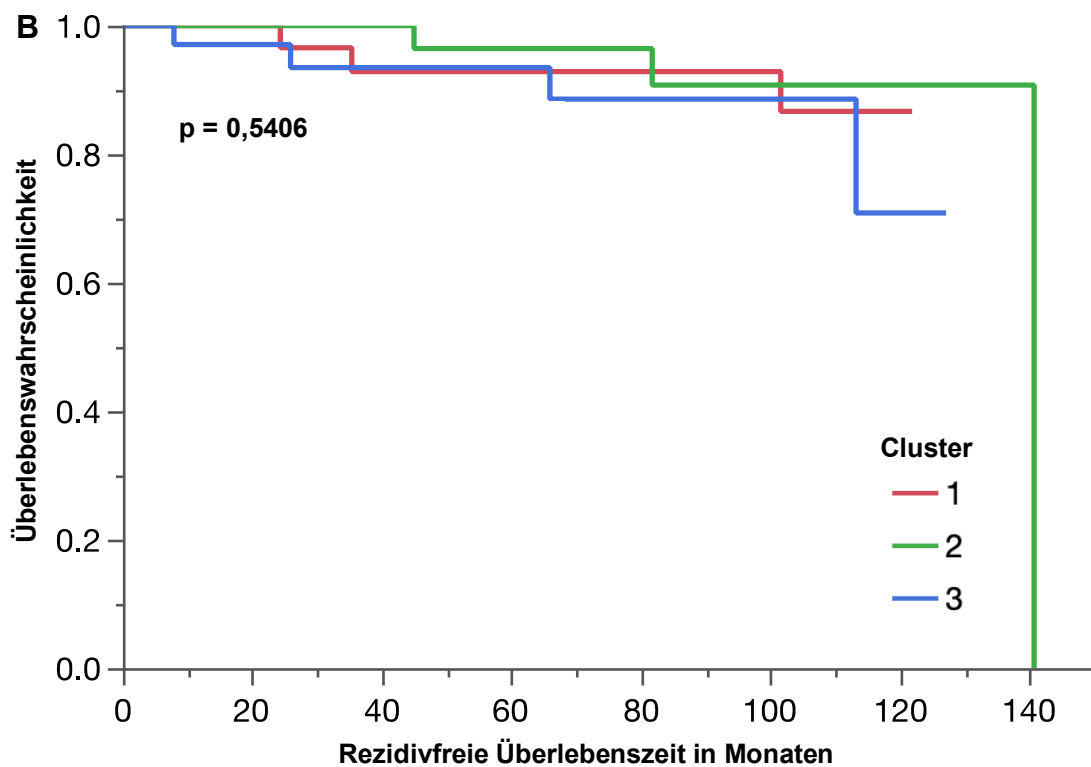
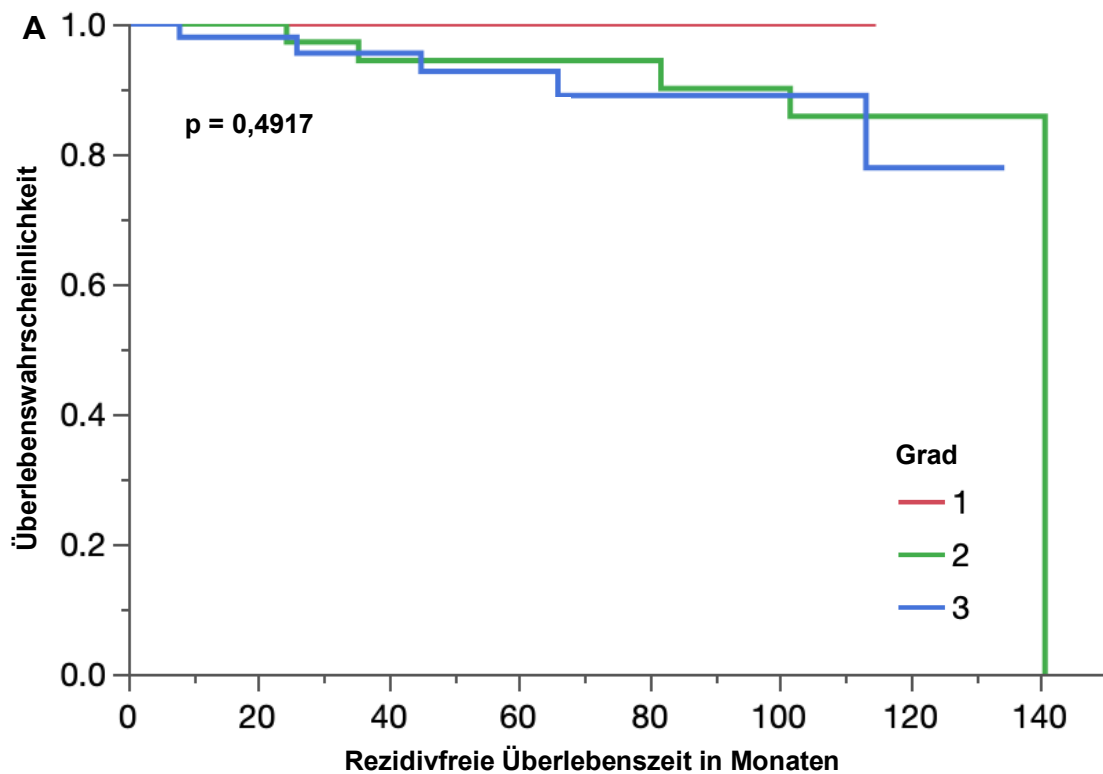
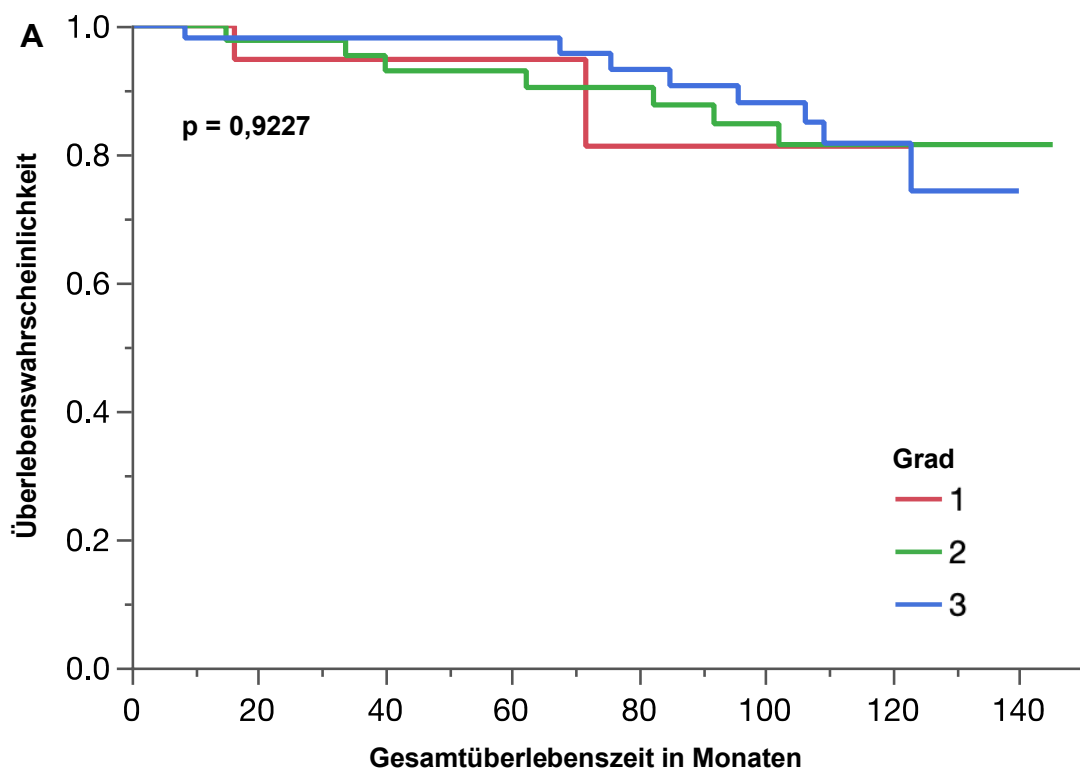


Abbildung 38: Kaplan-Meyer-Diagramme die Grading (A) und multiparametrische Cluster (B) in Relation zum Rezidivfreien Überleben setzen. Die Ergebnisse erweisen sich als nicht signifikant, und die kumulativen Inzidenzen der einzelnen Grade und Cluster variieren nur geringfügig.

3.2.5.2 Korrelation mit dem Gesamtüberleben (OS)

Ähnlich wie für das krankheitsfreie Überleben konnte im Rahmen dieser Studie weder für das herkömmliche Grading (p: 0,9227) noch für die multiparametrischen Cluster (p: 0,1212) ein signifikanter Zusammenhang mit dem Gesamtüberleben hergestellt werden. Die Kurven der Kaplan-Meyer-Diagramme zeigen keine deutliche Trennung. Nach 72 Monaten zeichnet sich ein leichter Anstieg der kumulativen Inzidenz der Fälle mit Grad 1 ab, doch nach 108 Monaten nähern sich die Kurven wieder einem ähnlichen Niveau an. Der p-Wert für die Korrelation der multiparametrischen Cluster mit dem Gesamtüberleben ist zwar niedriger, jedoch immer noch nicht signifikant. Interessanterweise weist das erste Cluster („low risk“) mit etwa 70% nach 100 Monaten die niedrigste Überlebenschance auf, während die Kurven von Cluster 2 und 3 einen weitestgehend ähnlichen Verlauf haben (siehe Abbildung 39).



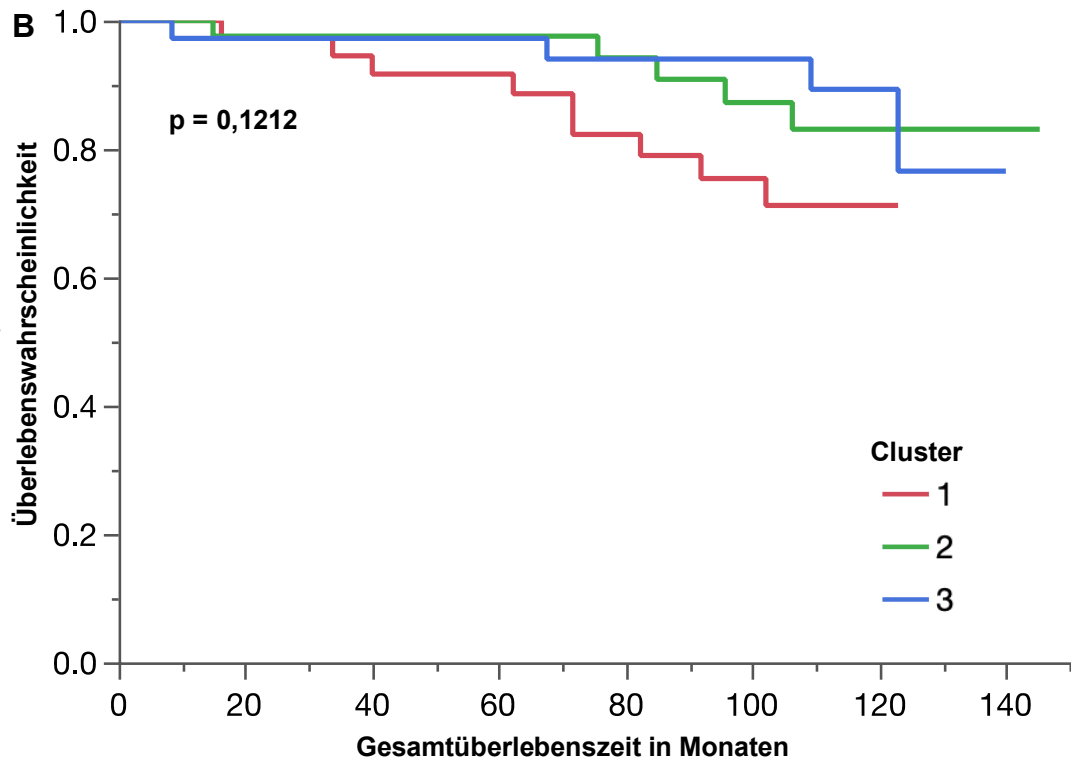


Abbildung 39: Kaplan-Meyer-Diagramme, die Grading (A) und multiparametrische Cluster (B) in Relation zum Gesamtüberleben setzen. Die Ergebnisse erweisen sich als nicht signifikant, und die kumulativen Inzidenzen der einzelnen Grade und Cluster variieren nur geringfügig. Eine Ausnahme stellt hierbei das „low risk“-Cluster dar, das eine höhere kumulative Inzidenz aufweist als die anderen Cluster.

4 Diskussion

Das Mammakarzinom stellt eine Tumorart dar, die infolge ihrer Prävalenz und der damit einhergehenden Belastung des Gesundheitssystems über mehrere Jahrzehnte hinweg einem ausgiebigen Forschungsinteresse unterliegt. Primäres Anliegen dieser Forschungsbemühungen ist die Verfeinerung des Tumorgradings, insbesondere auf molekularer Ebene, sowie die Implementierung digitaler Analyseverfahren mit dem Ziel, die Präzision und Standardisierung der pathologischen Beurteilung zu optimieren.

In diesem Zusammenhang nimmt die Untersuchung des Proliferationsmarkers Ki67 eine bedeutsame Rolle ein. Die IKWG konnte bei einem Cut-off-Wert von 20% für HR+ Tumoren einen signifikanten Zusammenhang mit einem schlechteren Gesamtüberleben feststellen (Robertson et al., 2020). Trotz dieser Evidenz existieren Einschränkungen hinsichtlich der Aussagekraft von Ki67. In Anbetracht hoher Inter- und Intraobserver-Variabilität kann besonders für Fälle, die sich im Grenzspektrum der Proliferationsrate (5-30%) befinden, keine definitive Aussage über die prognostische Relevanz des Markers getroffen werden (Shui et al., 2015). Dies stellt auch die IKWG in ihrer zuletzt herausgegebenen Empfehlung dar (Nielsen et al., 2021). Im klinischen Alltag erschwert dies die Entscheidungsfindung für therapeutische Maßnahmen, wie etwa die Applikation chemotherapeutischer Substanzen. Durch Normierung und Standardisierung wird angestrebt die Reproduzierbarkeit der Beurteilung zu steigern und somit eine optimierte personalisierte Therapie für die betroffenen Patientinnen zu ermöglichen.

Die Beurteilung von Mammakarzinomen anhand des Goldstandards wird von Patholog:innen als zeitaufwendige und mühsame Arbeit angesehen, die zudem durch die Subjektivität der evaluierenden Fachperson beeinflusst wird. Hierbei soll der Mitosemarker PhH3 Abhilfe schaffen (Puripat and Loharamtaweethong, 2019). Zusätzlich wurde kürzlich in aktuellen Studien ermittelt, dass dieser Marker als unabhängiger prädiktiver Faktor für das Gesamtüberleben und das krankheitsfreie Überleben angesehen werden kann (Ibrahim et al., 2023).

Sowohl bei der Bewertung von Ki67 als auch von PhH3 sollen digitale Analysemethoden, exemplarisch Softwarelösungen wie QuPath, dazu beitragen, die Subjektivität der pathologischen Beurteilung zu minimieren und somit eine homogenere Ergebnisfindung zu erreichen.

Darüber hinaus wird intensiv an multiparametrischen Ansätzen geforscht, die eine noch präzisere Klassifizierung von Mammakarzinomen ermöglichen sollen (Yu et al., 2021, Li et al., 2021).

Diese Studie zielte darauf ab, visuelle und digitale Auswertungsmethoden von Ki67 und PhH3 zu vergleichen und ihren praktischen Nutzen im klinischen Alltag zu evaluieren. Ferner sollte ermittelt werden, ob ein multiparametrischer Ansatz, der die Auswertung dieser beiden Marker einschließt, potenziell Einfluss auf den klinischen Verlauf der Erkrankung haben könnte.

4.1 Vergleich verschiedener Methoden zur Auswertung von Ki67 und PhH3

4.1.1 Visuelle Auswertung des Proliferationsindex (Ki67) – Schätzung vs. IKWG-Protokoll

Die Evaluation des Ki67-Proliferationsindex gründet bis dato auf der visuellen Beurteilung durch erfahrene Patholog:innen. Über einen längeren Zeitraum hinweg haben zahlreiche Einrichtungen Proliferationsraten gemäß individueller Protokolle analysiert, bis die IKWG eine einheitliche Methodik entwickelte, welche darauf abzielte, diesen Prozess zu standardisieren und die Interobserver-Variabilität weitestgehend zu reduzieren (Robertson et al., 2020). Im Rahmen dieser Studie wurde die lang etablierte Auswertungsmethode mittels globaler Einschätzung, die am Pathologischen Institut Tübingen über lange Zeit hinweg Verwendung fand, mit der Auswertung gemäß dem IKWG-Protokoll verglichen. Sowohl die 64 ausgewerteten Fälle der Entdeckungskohorte ($p < 0,001$, $R: 0,985$) als auch die 129 Resektate und die 125 Stanzbiopsien der Validierungskohorte (Resektate: $p < 0,001$, $R: 0,983$; Stanzbiopsien: $p < 0,001$, $R: 0,993$) zeigten eine bemerkenswert hohe Übereinstimmung zwischen den beiden Methoden. Daraus lässt sich schlussfolgern, dass die Ergebnisse mittels des IKWG-Protokolls die Schätzungen effektiv bestätigen und somit ein probates, standardisiertes Werkzeug zur Bestimmung des Ki67-Proliferationsindex darstellen, welches

jedoch wegen seiner besseren Interobserver-Reproduzierbarkeit der visuellen Schätzung im Alltag überlegen ist (Nielsen et al., 2021).

Es ist anzumerken, dass sämtliche Analysen von einer einzigen Histologin (Dr. Irene Gonzalez-Menendez) in zeitlichem Abstand durchgeführt wurden. Dies bedeutet einerseits, dass die Intraobserver-Variabilität zwischen den Methoden sehr gering zu sein scheint. Andererseits ist zu berücksichtigen, dass trotz der Erfahrung der auswertenden Histologin und ihrer weitreichenden Expertise in der Proliferationsraten-Bewertung, das Risiko eines Bestätigungsfehlers nicht gänzlich ausgeschlossen werden kann.

Da Interobserver-Variabilitäten bereits von vorherigen Studien dokumentiert wurden (Shui et al., 2015, Nielsen et al., 2021), empfiehlt sich für zukünftige Projekte eine Auswertung der Fälle durch diverse Patholog:innen, um eine fundiertere Aussage über den Nutzen des IKWG-Protokolls für die Reproduzierbarkeit der Resultate treffen zu können.

4.1.2 Visuelle Mitosezählung (PhH3 vs. HE)

Wie bereits erwähnt, ist die manuelle Auswertung von Mitosen im HE-gefärbten Gewebeschnitt bei Patholog:innen nicht sehr beliebt, da dieser Ansatz zeitaufwändig ist und ein hohes Maß an Aufmerksamkeit erfordert (Puripat and Loharamtaweethong, 2019). Vor diesem Hintergrund haben zahlreiche Forschende sich damit auseinandergesetzt, wie diese Verfahrensweise vereinfacht werden könnte. In vielen Studien stellt der Mitosemarker PhH3 die Lösung dar (van Steenhoven et al., 2020, Medani et al., 2021, Ibrahim et al., 2023). Dieser Marker erleichtert das Identifizieren und Auszählen von Mitosen erheblich und ermöglicht somit eine schnellere Analyse (Kim et al., 2017, Bosch et al., 2017). Innerhalb dieser Untersuchung konnte dieser subjektive Eindruck bestätigt werden. Aufgrund des ausgeprägten Kontrastes lassen sich Mitosefiguren eindeutiger erkennen und klarer von apoptotischen Zellen oder Artefakten abgrenzen.

Es fällt auf, dass verschiedene Arbeitsgruppen zu dem Konsens gelangen, dass unter Verwendung von PhH3 deutlich mehr Mitosen identifiziert werden können

als im Fall der HE-Färbung (Ibrahim et al., 2023, Mirzaian et al., 2020). Dieses Erkenntnis wurde in dieser Studie anhand der Stanzbiopsien der Entdeckungskohorte bestätigt. Im Vergleich zur HE-Färbung (M: 9, IQ: 10) wurden im PhH3-Schnitt (M: 43,8, IQ: 42,6) signifikant mehr Mitosen erfasst. Diese Veränderung zeigt sich auch in der Neubewertung der Fälle durch die am Grading orientierte Re-Klassifizierung auf Basis der PhH3-Färbung. Nach dem Re-Grading wurden 65,3% der Fälle als Grad 3 eingestuft, im Vergleich zu 77,8% nach der PhH3-Auswertung. Gleichwohl zeigen beide Methoden eine stark signifikante Korrelation ($p < 0,001$, $R: 0,675$). Zu diesem Schluss kamen unter anderem auch Jin et al. (Spearman correlation coefficient: 0,63, $p < 0.0001$) (Jin et al., 2019). Diese Resultate legen nahe, dass PhH3 den MAI von Mammakarzinomen adäquat widerspiegelt. Eine Bestätigung dieser Ergebnisse wurde in der Validierungskohorte erlangt (Korrelation HE und PhH3-Schnitt vom Resektat: $p: 0,0001$, $R: 0,78$). Auffällig ist hierbei, dass die Mitoseauswertung der Resektate sowohl im HE- (M: 8, IQ: 15) als auch im PhH3-Schnitt (M: 10, IQ: 15) sehr ähnlich ausfällt. Die Beurteilung der Mitosen erfolgte ebenfalls durch eine erfahrene Pathologin (Prof. Dr. Annette Staebler). Dennoch bleibt die Notwendigkeit, die Ergebnisse von anderen Patholog:innen zum Ausschluss von Bestätigungsfehlern erneut zu evaluieren.

Um die PhH3-Auswertung für den klinischen Alltag nutzbringend anzuwenden, bedarf es der Neu-Definition eines standardisierten Cut-Offs, welcher höher liegt als der etablierte Mitosescore nach Elston & Ellis (Elston and Ellis, 1991). In der Literatur werden unterschiedliche Werte genannt, etwa 13 Mitosen pro $1,59 \text{ mm}^2$ (Skaland et al., 2007), 13 Mitosen pro 2 mm^2 (van Steenhoven et al., 2020), oder 20 pro 3 mm^2 (Ibrahim et al., 2023). All diese Studien haben eine signifikante Korrelation zwischen einem hohen PhH3-Mitoseindex und einem schlechteren klinischen Verlauf nachweisen können.

4.1.2.1 Nachweis von falsch-negativen Mitosen

Die Präsenz von falsch-negativen Mitosefiguren wurde bereits von mehreren Forschungsgruppen dokumentiert (Tetzlaff et al., 2013, Ladstein et al., 2012). Im Kontext dieser Studie wiesen 19% der analysierten Fälle (Validierungskohorte)

solche falsch-negativen Mitosefiguren auf. Eine bemerkenswerte Beobachtung war hierbei, dass diese Erscheinung ausschließlich bei Resektaten auftrat. Insbesondere Fälle mit Grad 3 (n=19) und innerhalb dieser Klassifizierung vor allem solche mit dem TN-Subtyp (n=11) waren betroffen. Es ist auffällig, dass keine falsch-negativen Mitosen in den Stanzbiopsien festgestellt wurden, was darauf hinweist, dass eine rasche Fixation von Präparaten möglicherweise dazu beitragen kann, diesem Problem vorzubeugen. Zudem legt die Tatsache, dass überwiegend Grad 3-Tumoren betroffen waren nahe, dass es sich um stark proliferierende Tumoren handelt, die tendenziell größer sind und daher eine effiziente Fixierung erschweren. Die Schlussfolgerung wird ebenfalls durch die Erkenntnis gestützt, dass die meisten falsch-negativen Mitosen im Tumorkern lokalisiert waren. Dies stimmt auch mit den Vermutungen von Tetzlaff et al. überein, die eine mögliche Ursache in der präanalytischen Verarbeitung der Proben im Labor sehen (Tetzlaff et al., 2013).

Basierend auf der Studie von Ladstein et al., die keine signifikante Beeinflussung der Überlebensanalyse durch Berücksichtigung der nicht gefärbten Mitosen feststellen konnte, wurden diese in der vorliegenden Studie mit in die Mitosezählung einbezogen. Dennoch betonen auch Tetzlaff et al. in ihrer Arbeit die Bedeutung dieses Aspekts als wesentlichen limitierenden Faktor (Tetzlaff et al., 2013).

4.1.3 Visuelle vs. digitale Auswertung

Die Implementierung digitaler Auswertungsmethoden hat primär das Ziel, die pathologischen Befunde und Bewertungen zu standardisieren und zu objektivieren. Insbesondere zielt dieser Ansatz darauf ab, die Ergebnisse anspruchsvoll zu beurteilender Immunfärbungen, beispielsweise Ki67-Proliferationsindizes zwischen 5 und 30%, durch den Einsatz digitaler Methoden besser reproduzierbar zu gestalten. Darüber hinaus wird beabsichtigt, zeitaufwendige Verfahren, wie etwa die Auswertung von Mitosen, zu vereinfachen.

4.1.3.1 Auswertung des Proliferationsindex

Es existiert nur eine begrenzte Anzahl an Studien, welche unabhängig von der IKWG die digitale Auswertungsmethode mittels QuPath mit konventionellen visuellen Zählungen vergleichen. Die Literatur zu dieser Thematik zeigt uneinheitliche Einschätzungen hinsichtlich der Praktikabilität dieser Methode. Während Boyaci et al. den prognostischen Wert des IKWG-Protokolls zur digitalen Analyse bestätigen (Boyaci et al., 2021), verweisen Skjervold et al. darauf, dass die digitale Analyse von Ki67 mehr Fälle mit hohem Proliferationsindex identifiziert als die visuelle Methode. Die digitale Methode wird außerdem als unpraktikabel für den klinischen Alltag beschrieben (Skjervold et al., 2022). Es ist jedoch davon auszugehen, dass mit weiterer Optimierung der digitalen Methoden mit Anwendung von KI-basierten Protokollen, eine digitale, globale Bestimmung von Ki67 auch im Alltag Anwendung finden wird.

Im Rahmen der vorliegenden Studie konnte festgestellt werden, dass die digitale Methode in der Entdeckungskohorte sowohl im Vergleich zur Ki67-Auswertung mittels Schätzung als auch unter Anwendung des IKWG-Protokolls, eine starke und signifikante Korrelation aufweist ($p < 0,001$, $R: 0,884$). Diese Ergebnisse konnten ebenfalls durch die Validationskohorte bestätigt werden. Hierbei korrelierte die digitale Methode bei der Analyse der Resektate etwas schwächer mit den visuellen Methoden ($p < 0,001$, Schätzung: $0,879$; IKWG-Protokoll: $0,862$), als bei den Stanzbiopsien ($p < 0,001$, Schätzung: $R: 0,899$; IKWG-Protokoll: $R: 0,906$). Dies suggeriert, dass der Ki67-Proliferationsindex adäquat durch die digitale Methode mittels QuPath erfasst werden kann. Diese signifikanten Korrelationen lassen außerdem darauf schließen, dass die Ergebnisse der visuellen Schätzung und des IKWG-Protokolls mithilfe von QuPath von weniger erfahrenen Fachkräften erfolgreich reproduziert werden können. Eine solide Reproduzierbarkeit der digitalen Methode zwischen verschiedenen Forschenden wurde bereits in einer Studie von Acs et al. nachgewiesen (Acs et al., 2019).

Robertson et al. konstatierten, dass die digitale Methode (QuPath) im Median eine geringere Proliferationsrate aufweist im Vergleich zu visuellen Analysemethoden, die sich mit Auswertung von Hot Spots befassen (15,9% vs.

20%) (Robertson et al., 2020). Diese Tendenz konnte in der Entdeckungskohorte bestätigt werden. Anhand der Hot Spot Auswertung durch VM-Scope wurden hierbei im Median die höchsten Proliferationsraten bestimmt (M: 43% vs. IKWG: 37%; Schätzung: 35%; QuPath: 38,62%).

In Bezug auf die globale Auswertung der Entdeckungskohorte zeigte der Median der digitalen Analyse eine leichte Steigerung gegenüber der visuellen Auswertung mittels IKWG-Protokoll (38,62% vs. 37%). Die Validierungskohorte ermittelt für Resektate (QuPath: 5,85%; M: IKWG: 7,125%) und Stanzbiopsien (QuPath: 13,57%; M: IKWG: 15,75%,) durch die digitale Methode einen etwas niedrigeren Median als die visuellen Auswertungen.

Bemerkenswert ist, dass im Rahmen der globalen digitalen Analyse (QuPath) weniger HR+ HER2- Fälle als Luminal B klassifiziert werden, als mit den anderen verwendeten Methoden. In der Entdeckungskohorte werden durch die globale visuelle Methode mittels IKWG-Protokoll 20,3% und durch die Hot-Spot-Auswertung mittels VM-Scope 21,8% der Fälle als Luminal B Tumoren eingestuft. Die digitale Analyse mittels QuPath hingegen stuft nur 17,1% der Fälle als Luminal B ein. Diese Ergebnisse bestätigen sich in der Validierungskohorte. Sowohl die Auswertung der Resektate (17,8% vs. 20,1%) als auch der Stanzbiopsien (33,6% vs. 26,4%) zeigen im Vergleich zur globalen visuellen Methode (IKWG-Protokoll) eine geringe Zuordnung von Fällen zur Luminal B Subgruppe, wenn die Fälle mittels QuPath analysiert wurden.

Daraus ergibt sich folgende klinische Konsequenz: Würde man die prognostische Risikoeinschätzung der Stanze (und damit potenzielle Chemotherapieindikation) an den Surrogatmarkern für den molekularen Subtyp orientieren, würde die globale digitale Auswertung bei 7,2% aller Patientinnen ein geringeres Risiko angeben als die globale visuelle Schätzung bzw. die digitale Hot-Spot-Methode (VM-Scope). Bei diesen Patient:innen im Grenzbereich sollte in Erwägung gezogen werden, das Risiko zusätzlich zu überprüfen. Dazu gehören die Bestimmung des dynamischen Ki67 nach Einfluss einer kurzzeit-endokrinen Therapie sowie die Erstellung eines Genexpressionsprofils zur Bestimmung des molekularen Subtyps. Ähnliche Resultate erzielten Stalhammar et al., die

betonen, dass die digitale Auswertung deutlich besser mit der Genexpressionsanalyse durch PAM50 korreliert (76,6 % vs. 71,1 %) und somit unter Umständen geeigneter ist, um klinische Entscheidungen, wie das Verabreichen einer Chemotherapie zu treffen (Stalhammar et al., 2016). Um die Tragweite dieser Implikationen zu verstehen, sollten weitere Studien durchgeführt werden.

4.1.3.2 Mitosezählung

Die verfügbare wissenschaftliche Literatur zur digitalen Auswertung der Mitoserate ist begrenzt. Dennoch existieren Studien, welche visuelle und digitale Auswertungsmethoden vergleichen. In der Arbeit von Ibrahim et al. wurde eine hohe Übereinstimmung zwischen der visuellen und digitalen Auswertung von PhH3 beobachtet (Intraclass Korrelations Koeffizient (ICC): 0,824). Des Weiteren zeigten sie eine moderate Korrelation zwischen der digitalen PhH3-Analyse und der visuellen Mitosezählung am HE-Präparat (ICC: 0,405). Zu einer ähnlichen Schlussfolgerung gelangt auch diese Studie. In der Entdeckungskohorte konnte eine starke signifikante Korrelation sowohl zwischen der HE-Auswertung und der digitalen Analyse von PhH3-gefärbten Proben ($p < 0,001$, $R: 0,635$), als auch zwischen der visuellen und digitalen Analyse von PhH3-Präparaten ($p < 0,001$, $R: 0,860$) festgestellt werden. Letzteres konnte auch in der Validierungskohorte bestätigt werden. Hierbei korrelierten digitale und visuelle Auswertung von PhH3-gefärbten Stanzbiopsien ($p < 0,001$, $R: 0,806$) und Resektaten ($p < 0,001$, $R: 0,608$) ebenfalls stark miteinander.

Eine weitere Erkenntnis, die unter anderem in der Studie von Jin et al. dokumentiert wurde, ist, dass die digitale Analyse scheinbar mehr Mitosen auszählt als bei einer manuellen Auswertung (Jin et al., 2019). In der vorliegenden Entdeckungskohorte wurde eine ähnliche Tendenz beobachtet. Der Median der visuellen Auswertung (M: 43,8) lag deutlich unter dem der digitalen (M: 66,2). Ähnliche Ergebnisse zeichneten in der Validierungskohorte ab (Resektate: Visuell: M: 3, digital: M: 3,36; Stanzbiopsien Visuell: M: 3, digital: M: 9,4).

Die höheren Werte der digitalen PhH3 Auswertung in dieser Studie dürften größtenteils darauf zurückzuführen sein, dass nur ein geringer Anteil von Stromagewebe oder Artefakten in die Berechnung des MAI einbezogen wurde. Allerdings hat sich im Verlauf der Studie gezeigt, dass der angewandte Auswertungsansatz in der Praxis sehr viel Zeit kostet. Insbesondere die präzise Annotation einzelner Tumornester innerhalb der Resektatschnitte erwies sich als zeitaufwändig und nicht für den klinischen Alltag geeignet. Daher konnte nur eine begrenzte Anzahl von 45 Resektaten und 48 Stanzbiopsien aus der Validierungskohorte digital ausgewertet werden. Ein globaler Ansatz könnte voraussichtlich die mitotische Aktivität eines Tumors präziser darstellen. In diesem Kontext wäre eine automatische Erkennung der invasiven Tumorfläche durch eine KI sinnvoll. Ein ähnliches Projekt wird derzeit von einem weiteren PhD Doktoranden (Simon Holdenried-Krafft) im Rahmen dieses Graduiertenkollegs (GRK 2543) bearbeitet.

Für den klinische Einsatz empfiehlt sich daher zu diesem Zeitpunkt eher ein Ansatz, bei dem Hot-Spots an der Invasionsfront ausgewertet werden, wie er von verschiedenen Forschungsgruppen angewendet wurde (Jin et al., 2019, Ibrahim et al., 2023, Skaland et al., 2007). Wegen der höheren Einschätzung der Mitoserate nach Färbung mit PhH3 empfiehlt sich hier die Festlegung neuer Grenzwerte für den Mitosescore beim Grading.

4.2 Unterschied zwischen der Auswertung des Proliferations- und Mitoseindex von Stanzbiopsien und Resektaten

4.2.1 Auswertung des Proliferationsindex

Die vorherigen Ergebnisse veranschaulichen, dass die Proliferationsindizes der Resektate im Durchschnitt niedriger ausfallen als die der Stanzbiopsien (Schätzung: 20,05% vs. 12,79%, IKWG-Protokoll: 19,43% vs. 12,13%, QuPath: 20,23% vs. 11,12%). Die Korrelationen zwischen den verschiedenen Probenotypen sind bei allen angewandten Methoden moderat signifikant ($p < 0,001$, QuPath: $R: 0,41$, IKWG-Protokoll: $R: 0,40$, visuelle Schätzung: $R: 0,33$).

Da sämtliche Proben in domo angefertigt wurden, ist eine Diskrepanz aufgrund unterschiedlicher Aufbereitung im Labor weitestgehend auszuschließen. In Übereinstimmung mit Acs et al., die in ihrer Studie zu ähnlichen Ergebnissen gelangten, wird vermutet, dass die Ursache für diese Differenzen in der Zeit bis zur Fixierung liegen könnte. Im Durchschnitt sind Resektionspräparate einer längeren Periode warmer Ischämie ausgesetzt und werden erst später in Formalin eingelegt. Daher besteht die Möglichkeit, dass die Zellen in einer hypoxischen Umgebung mithilfe der verbliebenen intrazellulären Glucose bis zum G2/M-Checkpoint weiterhin Proliferation aufweisen und der Zellzyklus dann zum Stillstand kommt. Sobald dieser erreicht ist, weisen die Zellen keine Proliferation mehr auf und werden somit auch nicht mehr durch MIB-1 angefärbt (Lehr et al., 2013). Weiterhin wird angenommen, dass aufgrund der Größe der Resektatproben die Fixierung mit Formalin länger dauert, wodurch während dieser Phase ein Abbau der Epitope stattfinden kann. Im Gegensatz dazu werden Stanzbiopsien schneller entnommen und weisen aufgrund ihrer geringeren Größe eine kürzere Fixierungszeit auf (Acs et al., 2022, Lehr et al., 2013).

Die logische Konsequenz aus der niedrigeren Proliferationsrate bei Resektaten besteht darin, dass diese Proben im Gegensatz zu den Stanzbiopsien häufiger dem molekularen Subtyp Luminal A zugeordnet werden.

In Anbetracht dieser Erkenntnisse empfiehlt die IKWG, vorrangig Stanzbiopsien für die Ki67-Proliferationsanalyse zu verwenden, um die Verfälschung des Proliferationsindex durch Fixationsartefakte zu minimieren (Acs et al., 2019).

4.2.2 Mitosezählung

Ähnlich wie bei der Proliferationsanalyse zeigt sich im Rahmen dieser Studie, dass die Mitoserate der Stanzbiopsien anhand von PhH3 gefärbten Schnitten (M: 12,5,) im Vergleich zu den Resektaten (M: 10) etwas höher ausfällt. Zudem manifestiert sich eine stark signifikante Assoziation zwischen den verschiedenen Probenotypen ($p: 0,0001$, $R: 0,76$), was darauf hindeutet, dass ein erhöhter MAI sowohl im Resektat als auch in der Stanzbiopsie konsistent auftritt.

Diese Ergebnisse widersprechen der Annahme, dass bei Resektionspräparaten der MAI tendenziell höher ausfällt als bei Stanzbiopsien. Diese Erkenntnis wurde von mehreren Studien im Rahmen der Analyse von HE-Präparaten dokumentiert (Harris et al., 2003, Lehr et al., 2013). In einer Untersuchung von Lehr et al. wird dieser Umstand auf unklar definierte Artefakte zurückgeführt. Es wird vermutet, dass intraoperativer Stress und Trauma sowie die Einwirkung einer warmen Ischämie und eine länger andauernde Fixierungszeit in Formalin dazu führen können, dass die mitotische Aktivität in Resektionspräparaten durch allmählich abnehmende Energiereserven in der Metaphase am G2/M-Kontrollpunkt gestoppt wird. In dieser Phase können diese Mitosen von Patholog:innen leicht erkannt werden (Lehr et al., 2013). Im Gegensatz dazu werden Stanzbiopsien sehr rasch fixiert, wodurch die Zellteilungsprozesse frühzeitig unterbrochen werden. Dafür spricht auch die Ki67-Proliferationsrate in Resektaten im Vergleich zu Stanzbiopsien (Lehr et al., 2013).

Die Ergebnisse der vorliegenden Studie lassen vermuten, dass die von Lehr et al. erwähnten Artefakte in der PhH3-Immunhistochemie besser erkannt werden können und somit nicht in die Mitosezählung einfließen. Diese Annahme wird ebenfalls durch eine Studie von Kwok et al. gestützt, in der keine deutliche Diskrepanz in der der PhH3-Mitosezahl zwischen Stanzbiopsien (M: 13) und Resektaten (M:13) festgestellt wurde (Kwok et al., 2010). Wie bereits zuvor erwähnt, kann jedoch aufgrund der Auswertung durch eine einzige Pathologin ein Bestätigungsfehler nicht gänzlich ausgeschlossen werden. Daher sollten zukünftige Studien auch die Interobserver-Variabilität untersuchen.

4.3 Einfluss der Clusteranalyse auf die Tumorklassifizierung und den klinischen Verlauf

Mithilfe multiparametrischer Analyseansätze streben Forschende an, die Klassifizierung von Mammakarzinomen weiter zu präzisieren, um individuellere Therapiemöglichkeiten zu ermöglichen. Verschiedene Forschungsgruppen haben bereits eine robuste Korrelation mit dem rezidivfreien Überleben sowie dem Gesamtüberleben nachgewiesen (Li et al., 2021, Xiang et al., 2022, Yu et al., 2021). Zudem zeigen bereits entwickelte Analysesysteme eine gute

Prädiktion in Bezug auf das Grading und den molekularen Subtyp der Tumoren (Xi et al., 2021, Xu et al., 2022, Ohashi et al., 2023). In die multiparametrischen Klassifikationen fließen unter anderem Patientenalter, ethnische Herkunft, Hormonrezeptorstaus (ER, PR, HER2), Tumolvolumen, Lymphknotenstatus, Grading, histologischer Subtyp, Ki67-Index (Yu et al., 2021, Li et al., 2021), makro- und mikrokollagene Signaturen im Tumorgewebe (Xi et al., 2021), Tumordinfiltration und Karzinomstammzellpräsenz (Xiang et al., 2022) ein. Diese Parameter betreffen vor allem morphologische Eigenschaften. Nur wenige Studien schließen molekulare Faktoren wie den Ki67-Proliferationsindex mit ein (Yu et al., 2021).

Im Rahmen dieser Studie wurden neben den morphologischen Eigenschaften wie tubuläre Formation und nukleärer Pleomorphismus auch die visuelle Mitosezählung anhand von PhH3-gefärbten Schnitten und der digital ausgewertete Proliferationsindex betrachtet. Hohe Expressionsraten beider Marker korrelieren mit einem ungünstigeren klinischen Verlauf von Mammakarzinomen (Ibrahim et al., 2023, Nielsen et al., 2021). Zudem zeigte sich, dass die visuelle Analyse von PhH3 und die digitale Analyse von Ki67 sehr reproduzierbar sind (Ibrahim et al., 2023, Acs et al., 2019). In der Validierungskohorte wurden die relevanteren morphologischen Bewertungen des Resektates (tubuläre Formation und nukleärer Pleomorphismus) sowie die molekularen Auswertungen der zugehörigen Stanzbiopsien (PhH3, Ki67) verwendet. Letztere wurden ausgewählt, da sie laut Literaturrecherche die Tumorphiliferation besser repräsentieren als Resektate (Lehr et al., 2013, Acs et al., 2022).

Die Clusterbildung mittels der oben genannten Parameter führte sowohl in der Entdeckungskohorte als auch in der Validierungskohorte zu einer vielversprechenden Neugruppierung der Fälle in drei Subgruppen. Insbesondere Fälle mit Grad 2 und Grad 3 gliederten sich neu ein.

Leider konnte in der anschließenden Korrelation der Cluster mit den verfügbaren klinischen Daten keine Signifikanz für das Gesamtüberleben ($p: 0,1212$) oder das rezidivfreie Überleben ($p: 0,5406$) festgestellt werden. Gleiches gilt jedoch auch

für das herkömmliche Grading (Gesamtüberleben: $p: 0,9227$, rezidivfreies Überleben $p: 0,4917$). Dies legt nahe, dass diese Kohorte im Vergleich zu anderen multiparametrischen Studien sehr klein ist. Beispielsweise untersuchten Li et al. und Yu et al. jeweils mehr als 1000 Fälle (Li et al., 2021, Yu et al., 2021). Zudem wurde in der Validierungskohorte nicht das Alter der Patientinnen berücksichtigt. Dies ist vermutlich auch der Grund, warum im ersten Cluster, das nach morphologischen und molekularen Eigenschaften das geringste Risiko aufweisen sollte, tendenziell eine sehr geringe Überlebenschance sowohl für das Gesamtüberleben als auch für das rezidivfreie Überleben aufweist.

Angesichts der vielversprechenden Ergebnisse des Clusterings empfiehlt sich neben einer Vergrößerung der Kohorte auch eine Neustrukturierung derselben unter Berücksichtigung des Alters der Patientinnen. Eine sinnvolle Unterteilung könnte durch eine Altersgrenze von 50-55 Jahren erfolgen (Skaland et al., 2007, Yu et al., 2021). Durch diese Vorgehensweise ließe sich vermeiden, dass zu große biologische Unterschiede oder stark variierende Behandlungsstrategien die Ergebnisse beeinflussen. Darüber hinaus könnte die Auswertung einer größeren Kohorte mit längeren Follow-Up Daten und größeren homogenen Subgruppen mit adjuvanter Hormontherapie oder Chemotherapie, bessere Chancen auf signifikantere Ergebnisse haben.

4.4 Fazit

Diese Studie widmete sich dem Vergleich unterschiedlicher Analysemethoden, sowohl visuell als auch digital, für die Proliferationsmarker Ki67 und den Mitosemarker PhH3. Zusätzlich wurden diese IHC-Marker im Kontext einer multiparametrischen Analyse mit dem klinischen Verlauf (Gesamtüberleben und rezidivfreies Überleben) korreliert.

Die Auswertung von Ki67 ergab, dass das IKWG-Protokoll ein wertvolles Instrument zur Standardisierung der Proliferationsindex-Analyse darstellt. Die Reproduzierbarkeit wurde zuvor unter anderem durch Robertson et al. bestätigt (Robertson et al., 2020). Gleichzeitig zeigte sich, dass die digitale Analyse durch weniger erfahrene Forscher:innen die Ergebnisse der visuellen Auswertung durch erfahrene Patholog:innen gut reproduzieren kann. Dies legt nahe, dass Softwarelösungen wie QuPath, wenn sie in Abstimmung mit Protokollen der IKWG verwendet werden, dazu beitragen können, pathologische Befunde weiter zu objektivieren. Es wurde auch festgestellt, dass im Rahmen dieser Studie die Verwendung von digitaler Analyse (QuPath) dazu führt, dass weniger Fälle als Luminal B eingestuft werden. Dies könnte darauf hinweisen, dass die Proliferationsraten visuell tendenziell überschätzt werden und dies möglicherweise Auswirkungen auf die chemotherapeutische Behandlung haben könnte. Ein ähnliches Ergebnis erzielten auch Stalhammar et al.. Sie stellen klar, dass alle von ihnen verwendeten digitalen Auswertungsmethoden die manuellen in Bezug auf die Klassifikation von Luminal-B-Tumoren übertreffen. Dies belegen sie unter anderem durch eine höhere prozentuale Übereinstimmung der digitalen Methoden mit dem PAM50-Subtyp (76,6 % vs. 71,1 %) (Stalhammar et al., 2016).

Grundsätzlich wurden in dieser Studie die Proliferationsraten der Resektate im Durchschnitt niedriger als die der Stanzbiopsien eingestuft, was die Ergebnisse von Acs et al. bestätigt (Acs et al., 2022). Eine mögliche Ursache hierfür könnte der Abbau der Epitope aufgrund der längeren Exposition gegenüber warmer Ischämie und verzögerten Fixierung von Resektaten sein (Lehr et al., 2013, Acs et al., 2022)

In Bezug auf die Mitoseauswertung mittels PhH3 zeigte sich in der Entdeckungskohorte, dass mehr Mitosen im IHC-Schnitt von Stanzbiopsien im Vergleich zum HE-Präparat gefunden wurden. Dies ist konsistent mit Ergebnissen anderer Studien (Ibrahim et al., 2023, Skaland et al., 2007). Interessanterweise ergab sich, dass die Mitoserate in PhH3-Schnitten von Resektaten ähnlich der im HE-Schnitt war. Zudem wiesen die zugehörigen Stanzbiopsien leicht höhere Mitoseraten auf. Diese Befunde stehen im Gegensatz zu Forschungsergebnissen, bei denen im HE-Präparat die Mitoseraten von Resektaten höher als die von Stanzbiopsien ausfielen (Lehr et al., 2013, Harris et al., 2003). Eine mögliche Erklärung hierfür könnte sein, dass Artefakte in PhH3-Schnitten besser differenziert werden können und somit nicht in die Mitosezählung einfließen. Diese Hypothese wird durch die Ergebnisse der Studie von Kwok et al. gestützt, bei der die Mitoseraten von Resektaten und Stanzbiopsien in PhH3-Schnitten kaum Unterschiede aufwiesen (Kwok et al., 2010).

Die visuelle Auswertung von PhH3 erwies sich als praktikabel im klinischen Alltag, da Mitosefiguren deutlich schneller und effizienter differenziert werden können als im HE-Präparat. Diese Erkenntnis stimmt mit den Ergebnissen anderer Studien überein (Ibrahim et al., 2023, Mirzaian et al., 2020, Skaland et al., 2007). Nichtsdestotrotz sollte bei der Auswertung von PhH3 beachtet werden, dass das Auftreten von falsch-negativen Mitosen ein limitierender Faktor sein kann (Tetzlaff et al., 2013). Diese nicht reaktiven Mitosefiguren wurden auch in dieser Studie nachgewiesen, besonders bei TN-Fällen mit Grad 3. Eine mögliche Erklärung dafür könnte eine unzureichende Fixierung im Tumorzentrum sein, da in den betroffenen Fällen gefärbte Mitosefiguren eher in den äußeren Zonen nachgewiesen werden können.

Die digitale globale Auswertungsmethode, besonders im Rahmen der Analyse großer Tumorproben, erwies sich als eher unpraktikabel. Hierbei wäre eine automatische Erkennung der invasiven Tumorfläche, wie sie von Simon Holdenried-Krafft im Rahmen des GRK 2543 entwickelt wird, von Nutzen. Für einen klinischen Einsatz empfiehlt sich daher eher eine Methode, die sich auf

Hot-Spots im Bereich der Invasionsfront konzentriert (Jin et al., 2019, Ibrahim et al., 2023, Skaland et al., 2007).

Um Bestätigungsfehler bei den visuellen Auswertungen von Ki67 und PhH3 durch eine einzige Person auszuschließen, sollten weitere Expert:innen diese Ergebnisse erneut analysieren. Dies würde auch zur Bewertung der Reproduzierbarkeit beitragen.

Im Rahmen dieser Studie konnte keine signifikante Korrelation der multiparametrischen Cluster mit den Überlebensdaten der Patientinnen festgestellt werden. Dies könnte mit der fehlenden Aufteilung der Kohorte nach Alter und der im Vergleich zu anderen Studien geringen Fallzahl zusammenhängen. Dennoch sollten diese Ergebnisse mit optimierten Rahmenbedingungen erneut überprüft werden, da das Clustering eine vielversprechende Neugruppierung von G2- und G3-Fällen sowohl in der Entdeckungs- als auch in der Validierungskohorte zeigt.

5 Zusammenfassung

Das Mammakarzinom ist die häufigste Tumorerkrankung bei Frauen in Deutschland. Angesichts der wachsenden Bedeutung personalisierter Medizin und der Belastung des Gesundheitssystems besteht ein zunehmendes Bedürfnis nach präziseren und objektiveren Diagnostikmethoden. Konventionelle Ansätze zur Bewertung der Proliferation, wie der Ki67-Proliferationsindex und die Mitosezählung, sind durch Interobserver-Variabilitäten und der daraus resultierenden begrenzten Reproduzierbarkeit eingeschränkt. Besonders betroffen sind Grad-2-Tumoren mit HR+, HER2-negativem Status, bei denen der Ki67-Index die Wahl der Therapie (Chemotherapie oder Hormontherapie) maßgeblich beeinflusst.

In dieser Arbeit wurden visuelle und digitale Analysemethoden zur Bewertung der Proliferationsrate (Ki67) und der Mitosezählung (PhH3) miteinander verglichen. Zusätzlich wurden morphologische und immunhistochemische Parameter in einem multiparametrischen Cluster kombiniert, um die Auswirkungen auf das Gesamtüberleben sowie das rezidivfreie Überleben zu evaluieren und mit dem herkömmlichen Grading zu vergleichen. Es wurden 72 Fälle in einer Entdeckungskohorte und 131 Fälle in einer Validierungskohorte analysiert. Die digitale Analyse erfolgte mithilfe der Software QuPath, während die visuelle Bewertung nach etablierten Protokollen durchgeführt wurde.

Die Ergebnisse zeigen eine signifikante Korrelation zwischen der digitalen und visuellen Bewertung des Ki67-Index. Interessanterweise wurden mit digitaler Analyse weniger Tumoren der molekularen Subgruppe Luminal B zugeordnet, was klinische Therapieentscheidungen beeinflussen könnte. Fixierungsartefakte führten bei Resektaten zu einer geringeren Proliferationsrate und Mitoseaktivität im Vergleich zu Stanzbiopsien. Zudem zeigte die digitale Auswertung des PhH3-Index eine höhere Sensitivität bei der Mitosezählung als die Analyse von HE-Präparaten. Der multiparametrische Ansatz erbrachte keine signifikanten Verbesserungen, führte jedoch zu einer Neugruppierung von Fällen der Grade 2 und 3 in der Clusteranalyse.

Diese Ergebnisse unterstreichen das Potenzial digitaler Methoden zur Verbesserung der diagnostischen Präzision und Objektivität bei Mammakarzinomen. Die klinische Relevanz der digitalen Subgruppierung erfordert jedoch weitere Untersuchungen, insbesondere im Hinblick auf ihre Auswirkungen auf die Therapieentscheidungen.

6 Literaturverzeichnis

- ACS, B., LEUNG, S. C. Y., KIDWELL, K. M., ARUN, I., AUGULIS, R., BADVE, S. S., BAI, Y., BANE, A. L., BARTLETT, J. M. S., BAYANI, J., BIGRAS, G., BLANK, A., BUIKEMA, H., CHANG, M. C., DIETZ, R. L., DODSON, A., FINEBERG, S., FOCKE, C. M., GAO, D., GOWN, A. M., GUTIERREZ, C., HARTMAN, J., KOS, Z., LAENKHOLM, A. V., LAURINAVICIUS, A., LEVENSON, R. M., MAHBOUBI-ARDAKANI, R., MASTROPASQUA, M. G., NOFECH-MOZES, S., OSBORNE, C. K., PENNAULT-LLORCA, F. M., PIPER, T., QUINTAYO, M. A., RAU, T. T., REINHARD, S., ROBERTSON, S., SALGADO, R., SUGIE, T., VAN DER VEGT, B., VIALE, G., ZABAGLO, L. A., HAYES, D. F., DOWSETT, M., NIELSEN, T. O., RIMM, D. L., INTERNATIONAL KI67 IN BREAST CANCER WORKING GROUP OF THE BREAST INTERNATIONAL, G. & NORTH AMERICAN BREAST CANCER, G. 2022. Systematically higher Ki67 scores on core biopsy samples compared to corresponding resection specimen in breast cancer: a multi-operator and multi-institutional study. *Modern Pathology* 10: 1362-1369
- ACS, B., PELEKANOU, V., BAI, Y., MARTINEZ-MORILLA, S., TOKI, M., LEUNG, S. C. Y., NIELSEN, T. O. & RIMM, D. L. 2019. Ki67 reproducibility using digital image analysis: an inter-platform and inter-operator study. *Laboratory Investigation* 99: 107-117.
- AEFFNER, F., ZARELLA, M. D., BUCHBINDER, N., BUI, M. M., GOODMAN, M. R., HARTMAN, D. J., LUJAN, G. M., MOLANI, M. A., PARWANI, A. V., LILLARD, K., TURNER, O. C., VEMURI, V. N. P., YUIL-VALDES, A. G. & BOWMAN, D. 2019. Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. *Journal of Pathology Informatics* 10: 9.
- AYSOLA, K., DESAI, A., WELCH, C., XU, J., QIN, Y., REDDY, V., MATTHEWS, R., OWENS, C., OKOLI, J., BEECH, D. J., PIYATHILAKE, C. J., REDDY, S. P. & RAO, V. N. 2013. Triple Negative Breast Cancer - An Overview. *Hereditary Genetics* 2013: Suppl 2: 001
- BANKHEAD, P. 2022. Developing image analysis methods for digital pathology. *The Journal of Pathology* 257: 391-402.
- BANKHEAD, P., LOUGHREY, M. B., FERNANDEZ, J. A., DOMBROWSKI, Y., MCART, D. G., DUNNE, P. D., MCQUAID, S., GRAY, R. T., MURRAY, L. J., COLEMAN, H. G., JAMES, J. A., SALTO-TELLEZ, M. & HAMILTON, P. W. 2017. QuPath: Open source software for digital pathology image analysis. *Scientific Reports* 7: 16878.
- BARISONI, L., LAFATA, K. J., HEWITT, S. M., MADABHUSHI, A. & BALIS, U. G. J. 2020. Digital pathology and computational image analysis in nephropathology. *Nature Reviews Nephrology* 16: 669-685.
- BAXI, V., EDWARDS, R., MONTALTO, M. & SAHA, S. 2022. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Modern Pathology* 35: 23-32.
- BLOOM, H. J. & RICHARDSON, W. W. 1957. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *British Journal of Cancer* 11: 359-377.

- BOEKHOUT, A. H., BEIJNEN, J. H. & SCHELLENS, J. H. 2011. Trastuzumab. *Oncologist* 16: 800-810.
- BÖGER, C., BEHRENS, H., WARNEKE, V. & RÖCKEN, C. 2015. Ki67-Proliferationsindex beim Magenkarzinom – Putativer Marker für intratumorale Heterogenität, als Prognoseindikator jedoch ungeeignet. *Zeitschrift für Gastroenterologie* 53: KG015.
- BOSCH, D. E., KILGORE, M. R., SCHMIDT, R. A., SWANSON, P. E., RENDI, M. H. & CHANG, O. H. 2017. Comparison of Proliferation Markers Ki67 and Phosphohistone-H3 (pHH3) in Breast Ductal Carcinoma In Situ. *Applied Immunohistochemistry and Molecular Morphology* 25: 543-547.
- BOYACI, C., SUN, W., ROBERTSON, S., ACS, B. & HARTMAN, J. 2021. Independent Clinical Validation of the Automated Ki67 Scoring Guideline from the International Ki67 in Breast Cancer Working Group. *Biomolecules* 11: 1612.
- CATTORETTI, G., BECKER, M. H., KEY, G., DUCHROW, M., SCHLÜTER, C., GALLE, J. & GERDES, J. 1992. Monoclonal antibodies against recombinant parts of the Ki-67 antigen (MIB 1 and MIB 3) detect proliferating cells in microwave-processed formalin-fixed paraffin sections. *Journal of Pathology* 168: 357-363.
- DAHLROT, R. H., BANGSØ, J. A., PETERSEN, J. K., ROSAGER, A. M., SØRENSEN, M. D., REIFENBERGER, G., HANSEN, S. & KRISTENSEN, B. W. 2021. Prognostic role of Ki-67 in glioblastomas excluding contribution from non-neoplastic cells. *Scientific Reports* 11: 17918.
- DANNEHL, D. & HARTKOPF, A. D. 2022. Personalisierte Medizin – innovative Therapiemöglichkeiten für das metastasierte Mammakarzinom. *Der Gynäkologe* 55: 335-343.
- DE AZAMBUJA, E., CARDOSO, F., DE CASTRO, G., JR., COLOZZA, M., MANO, M. S., DURBECQ, V., SOTIRIOU, C., LARSIMONT, D., PICCART-GEBHART, M. J. & PAESMANS, M. 2007. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. *British Journal of Cancer* 96: 1504-13.
- DE, M. R. J. F., DE MEDEIROS, R. S. S., BRAGHIROLI, M. I., GALVÃO, B., NETO, J. E. B., MUNHOZ, R. R., GUERRA, J., NONOGAKI, S., KIMURA, L., PFIFFER, T. E., DE CASTRO, G., JR., HOFF, P. M., FILHO, D. R., COSTA, F. P. & RIECHELMANN, R. P. 2017. Expression of ERCC1, Bcl-2, Lin28a, and Ki-67 as biomarkers of response to first-line platinum-based chemotherapy in patients with high-grade extrapulmonary neuroendocrine carcinomas or small cell lung cancer. *Ecancermedicalscience* 11: 767.
- DESSAUVAGIE, B. F., THOMAS, C., ROBINSON, C., FROST, F. A., HARVEY, J. & STERRETT, G. F. 2015. Validation of mitosis counting by automated phosphohistone H3 (PHH3) digital image analysis in a breast carcinoma tissue microarray. *Pathology* 47: 329-334.
- DOWSETT, M., NIELSEN, T. O., A'HERN, R., BARTLETT, J., COOMBES, R. C., CUZICK, J., ELLIS, M., HENRY, N. L., HUGH, J. C., LIVELY, T., MCSHANE, L., PAIK, S., PENAULT-LLOORCA, F., PRUDKIN, L., REGAN, M., SALTER, J., SOTIRIOU, C., SMITH, I. E., VIALE, G., ZUJEWSKI, J. A., HAYES, D. F. & INTERNATIONAL KI-67 IN BREAST CANCER WORKING, G. 2011. Assessment of Ki67 in breast cancer:

- recommendations from the International Ki67 in Breast Cancer working group. *Journal of the National Cancer Institute* 103: 1656-1664.
- ELDER, E., XU, D., HÖÖG, A., ENBERG, U., HOU, M., PISA, P., GRUBER, A., LARSSON, C. & BÄCKDAHL, M. 2003. KI-67 AND hTERT Expression Can Aid in the Distinction between Malignant and Benign Pheochromocytoma and Paraganglioma. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 16: 246-255.
- ELLIS, I. O. C., P.; HALES, S.; LEE, A. H. S.; PINDER, S. E.; RAKHA, E.; STEPHENSON, T.; AL-SAM, S.; DEB, R.; HANBY, A.; LIEBMANN, R.; PROVENZANO, E.; ROWLANDS, D.; WELLS, A.; ANDERSON, N.; GIRLING, A.; IBRAHIM, M.; MALLON, E.; QUINN, C.; 2016. Pathology reporting of breast disease in surgical excision specimens incorporating the dataset for histological reporting of breast cancer. *The Royal College of Pathologists G 148 LR.*: 1-160.
- ELSTON, C. W. & ELLIS, I. O. 1991. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 19: 403-410.
- GALEA, M. H., BLAMEY, R. W., ELSTON, C. E. & ELLIS, I. O. 1992. The Nottingham prognostic index in primary breast cancer. *Breast Cancer Research and Treatment* 22: 207-219.
- GERDES, J., LI, L., SCHLUETER, C., DUCHROW, M., WOHLLENBERG, C., GERLACH, C., STAHRM, I., KLOTH, S., BRANDT, E. & FLAD, H. D. 1991. Immunobiochemical and molecular biologic characterization of the cell proliferation-associated nuclear antigen that is defined by monoclonal antibody Ki-67. *American Journal of Pathology* 138: 867-873.
- GERDES, J., SCHWAB, U., LEMKE, H. & STEIN, H. 1983. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *International Journal of Cancer* 31: 13-20.
- GOLDHIRSCH, A., WINER, E. P., COATES, A. S., GELBER, R. D., PICCART-GEBHART, M., THURLIMANN, B., SENN, H. J. & PANEL, M. 2013. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Annals of Oncology* 24: 2206-2223.
- GUARNERI, V. & CONTE, P. 2009. Metastatic Breast Cancer: Therapeutic Options According to Molecular Subtypes and Prior Adjuvant Therapy. *The Oncologist* 14: 645-656.
- GURLEY, L. R., D'ANNA, J. A., BARHAM, S. S., DEAVEN, L. L. & TOBEY, R. A. 1978. Histone Phosphorylation and Chromatin Structure during Mitosis in Chinese Hamster Cells. *European Journal of Biochemistry* 84: 1-15.
- HALE, C. S., QIAN, M., MA, M. W., SCANLON, P., BERMAN, R. S., SHAPIRO, R. L., PAVLICK, A. C., SHAO, Y., POLSKY, D., OSMAN, I. & DARVISHIAN, F. 2013. Mitotic rate in melanoma: prognostic value of immunostaining and computer-assisted image analysis. *American Journal of Surgical Pathology* 37: 882-889.
- HAO Q, D. C., DENG Y, XU P, TIAN T, LIN S, WANG M, LIU K, SONG D, WU Y, GUO Y, DAI Z. 2018. Pooling analysis on prognostic value of PHH3

expression in cancer patients. *Cancer Management and Research* 10: 2279-2288

- HARRIS, G. C., DENLEY, H. E., PINDER, S. E., LEE, A. H. S., ELLIS, I. O., ELSTON, C. W. & EVANS, A. 2003. Correlation of Histologic Prognostic Factors in Core Biopsies and Therapeutic Excisions of Invasive Breast Carcinoma. *American Journal of Surgical Pathology* 27: 11-15.
- HUMPHRIES, M. P., MAXWELL, P. & SALTO-TELLEZ, M. 2021. QuPath: The global impact of an open source digital pathology system. *Computational and Structural Biotechnology Journal* 19: 852-859.
- I. A. MONTES MOJARRO, M. G., T. SCHOTTE, D. SILIMON, L. M. SERNA HIGUITA, HOLDENRIED-KRAFFT S., L. L. VOLMER, H. LENSCH, A. & FEND, S. A. F. 2023. Comparative analysis of morphological and immunophenotypic parameters of breast cancer NST using Artificial Intelligence. In: TÜBINGEN, U. (ed.) Poster ed. Tübingen.
- IBRAHIM, A., TOSS, M. S., MAKHLOUF, S., MILIGY, I. M., MINHAS, F. & RAKHA, E. A. 2023. Improving mitotic cell counting accuracy and efficiency using phosphohistone-H3 (PHH3) antibody counterstained with haematoxylin and eosin as part of breast cancer grading. *Histopathology* 82: 393-406.
- JIANG, Y., YANG, M., WANG, S., LI, X. & SUN, Y. 2020. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Communications* 40: 154-166.
- JIN, C., HUANG, Y., NASIM, M., YANG, Y. & LEE, L. 2019. Gastrointestinal Stromal Tumors Risk Stratification Utilizing Phospho-Histone H3 Evaluated by Manual Counting and Computer-Assisted Image Analysis. *International Journal of Surgical Pathology* 27: 706-712.
- JIN, X. & MU, P. 2015. Targeting Breast Cancer Metastasis. *Breast Cancer* 9: 23-34.
- KATALINIC, A., EISEMANN, N., KRAYWINKEL, K., NOFTZ, M. R. & HUBNER, J. 2020. Breast cancer incidence and mortality before and after implementation of the German mammography screening program. *International Journal of Cancer* 147: 709-718.
- KIM, J.-Y., JEONG, H. S., CHUNG, T., KIM, M., LEE, J. H., JUNG, W. H. & KOO, J. S. 2017. The value of phosphohistone H3 as a proliferation marker for evaluating invasive breast cancers: A comparative study with Ki67. *Oncotarget* 8: 65064-65076
- KLINTMAN, M., STRAND, C., AHLIN, C., BEGLERBEGOVIC, S., FJÄLLSKOG, M. L., GRABAU, D., GUDLAUGSSON, E., JANSSEN, E. A., LÖVGREN, K., SKALAND, I., BENDAHL, P. O., MALMSTRÖM, P., BAAK, J. P. & FERNÖ, M. 2013. The prognostic value of mitotic activity index (MAI), phosphohistone H3 (PPH3), cyclin B1, cyclin A, and Ki67, alone and in combinations, in node-negative premenopausal breast cancer. *PLoS One* 8: e81902.
- KOO, M. M., VON WAGNER, C., ABEL, G. A., MCPHAIL, S., RUBIN, G. P. & LYRATZOPOULOS, G. 2017. Typical and atypical presenting symptoms of breast cancer and their associations with diagnostic intervals: Evidence from a national audit of cancer diagnosis. *Cancer Epidemiology* 48: 140-146.

- KWOK, T. C., RAKHA, E. A., LEE, A. H., GRAINGE, M., GREEN, A. R., ELLIS, I. O. & POWE, D. G. 2010. Histological grading of breast cancer on needle core biopsy: the role of immunohistochemical assessment of proliferation. *Histopathology* 57: 212-219.
- LADSTEIN, R. G., BACHMANN, I. M., STRAUME, O. & AKSLEN, L. A. 2012. Prognostic importance of the mitotic marker phosphohistone H3 in cutaneous nodular melanoma. *Journal of Investigative Dermatology* 132: 1247-1252.
- LEHR, H.-A., ROCHAT, C., SCHAPER, C., NOBILE, A., SHANOUDA, S., VIJGEN, S., GAUTHIER, A., OBERMANN, E., LEUBA, S., SCHMIDT, M., C, C. R., DELALOYE, J.-F., SIMIANTONAKI, N. & SCHAEFER, S. C. 2013. Mitotic figure counts are significantly overestimated in resection specimens of invasive breast carcinomas. *Modern Pathology* 26: 336-342.
- LEITLINIENPROGRAMM ONKOLOGIE (DEUTSCHE KREBSGESELLSCHAFT, D. K. & AWMF) 2021. S3-Leitlinie Früherkennung, Diagnose, Therapie und Nachsorge des Mammakarzinoms, Version 4.3, AWMF Registernummer: 032-045OL, <http://www.leitlinienprogramm-onkologie.de/leitlinien/mammakarzinom/> [Zugriff 19.07.2023].
- LEUNG, S. C. Y., NIELSEN, T. O., ZABAGLO, L. A., ARUN, I., BADVE, S. S., BANE, A. L., BARTLETT, J. M. S., BORGQUIST, S., CHANG, M. C., DODSON, A., EHINGER, A., FINEBERG, S., FOCKE, C. M., GAO, D., GOWN, A. M., GUTIERREZ, C., HUGH, J. C., KOS, Z., LAENKHOLM, A. V., MASTROPASQUA, M. G., MORIYA, T., NOFECH-MOZES, S., OSBORNE, C. K., PENAUULT-LLORCA, F. M., PIPER, T., SAKATANI, T., SALGADO, R., STARCZYNSKI, J., SUGIE, T., VAN DER VEGT, B., VIALE, G., HAYES, D. F., MCSHANE, L. M., DOWSETT, M., INTERNATIONAL KI67 IN BREAST CANCER WORKING GROUP OF THE BREAST INTERNATIONAL, G. & NORTH AMERICAN BREAST CANCER, G. 2019. Analytical validation of a standardised scoring protocol for Ki67 immunohistochemistry on breast cancer excision whole sections: an international multicentre collaboration. *Histopathology* 75: 225-235.
- LI, N., FENG, L. W., LI, Z. N., WANG, J. & YANG, L. 2021. Construction and Validation of a Nomogram to Predict Overall Survival in Very Young Female Patients with Curatively Resected Breast Cancer. *Cancer Management and Research* 13: 6181-6190.
- LI, Z. H., HU, P. H., TU, J. H. & YU, N. S. 2016. Luminal B breast cancer: patterns of recurrence and clinical outcome. *Oncotarget* 7: 65024-65033.
- LIANG, Q., MA, D., GAO, R. F. & YU, K. D. 2020. Effect of Ki-67 Expression Levels and Histological Grade on Breast Cancer Early Relapse in Patients with Different Immunohistochemical-based Subtypes. *Scientific Reports* 10: 7648.
- ŁUKASIEWICZ, S., CZECZELEWSKI, M., FORMA, A., BAJ, J., SITARZ, R. & STANISŁAWEK, A. 2021. Breast Cancer-Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies-An Updated Review. *Cancers* 13: 4287.
- MAKKI, J. 2015. Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. *Clinical Medicine Insights: Pathology* 8: 23-31.

- MEDANI, H., ELSHIEKH, M. & NARESH, K. N. 2021. Improving precise counting of mitotic cells in mantle cell lymphoma using phosphohistone H3 (PHH3) antibody. *Journal of Clinical Pathology* 74: 646-649.
- MIKAMI, Y., UENO, T., YOSHIMURA, K., TSUDA, H., KUROSUMI, M., MASUDA, S., HORII, R., TOI, M. & SASANO, H. 2013. Interobserver concordance of Ki67 labeling index in breast cancer: Japan Breast Cancer Research Group Ki67 ring study. *Cancer Science* 104: 1539-1543.
- MIRZAIAN, E., TABATABAEI GHODS, Z. S., TAVANGAR, S. M., EMAMI, B., ORAIE, M., SAFYARI, R. & SAFFAR, H. 2020. Utility of PHH3 in Evaluation of Mitotic Index in Breast Carcinoma and Impact on Tumor Grade. *Asian Pacific Journal of Cancer Prevention* 21: 63-66.
- MITRI, Z., CONSTANTINE, T. & O'REGAN, R. 2012. The HER2 Receptor in Breast Cancer: Pathophysiology, Clinical Use, and New Advances in Therapy. *Chemother Research and Practice* 2012: 743193.
- MOO, T. A., SANFORD, R., DANG, C. & MORROW, M. 2018. Overview of Breast Cancer Therapy. *PET Clinics* 13: 339-354.
- NAHEED, S., HOLDEN, C., TANNO, L., PATTINI, L., PEARCE, N. W., GREEN, B., JAYNES, E., CAVE, J., OTTENSMEIER, C. H. & PELOSI, G. 2022. Utility of Ki-67 as a prognostic biomarker in pulmonary neuroendocrine neoplasms: a systematic review and meta-analysis. *BMJ Open* 12: e041961.
- NASR, M. R. & EL-ZAMMAR, O. 2008. Comparison of pHH3, Ki-67, and survivin immunoreactivity in benign and malignant melanocytic lesions. *American Journal of Dermatopathology* 30: 117-122.
- NIELSEN, P. S., RIBER-HANSEN, R., SCHMIDT, H. & STEINICHE, T. 2016. Automated quantification of proliferation with automated hot-spot selection in phosphohistone H3/MART1 dual-stained stage I/II melanoma. *Diagnostic Pathology* 11: 35.
- NIELSEN, T. O., LEUNG, S. C. Y., RIMM, D. L., DODSON, A., ACS, B., BADVE, S., DENKERT, C., ELLIS, M. J., FINEBERG, S., FLOWERS, M., KREIPE, H. H., LAENKHOLM, A. V., PAN, H., PENAULT-LLORCA, F. M., POLLEY, M. Y., SALGADO, R., SMITH, I. E., SUGIE, T., BARTLETT, J. M. S., MCSHANE, L. M., DOWSETT, M. & HAYES, D. F. 2021. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *Journal of the National Cancer Institute* 113: 808-819.
- OHASHI, A., KATAOKA, M., IIMA, M., HONDA, M., OTA, R., URUSHIBATA, Y., DOMINIK NICKEL, M., TOI, M., ZACKRISSON, S. & NAKAMOTO, Y. 2023. A multiparametric approach to predict triple-negative breast cancer including parameters derived from ultrafast dynamic contrast-enhanced MRI. *European Radiology* 33: 8132-8141.
- PELL, R., OIEN, K., ROBINSON, M., PITMAN, H., RAJPOOT, N., RITTSCHER, J., SNEAD, D. & VERRILL, C. 2019. The use of digital pathology and image analysis in clinical trials. *The Journal of Pathology: Clinical Research* 5: 81-90.
- PEROU, C. M., SORLIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., REES, C. A., POLLACK, J. R., ROSS, D. T., JOHNSEN, H., AKSLEN, L. A., FLUGE, O., PERGAMENSCHIKOV, A., WILLIAMS, C., ZHU, S. X.,

- LONNING, P. E., BORRESEN-DALE, A. L., BROWN, P. O. & BOTSTEIN, D. 2000. Molecular portraits of human breast tumours. *Nature* 406: 747-752.
- PURIPAT, N. & LOHARAMTAWEEHONG, K. 2019. Phosphohistone H3 (PHH3) as a surrogate of mitotic figure count for grading in meningiomas: a comparison of PHH3 (S10) versus PHH3 (S28) antibodies. *Virchows Archiv* 474: 87-96.
- RAKHA, E. A., REIS-FILHO, J. S., BAEHNER, F., DABBS, D. J., DECKER, T., EUSEBI, V., FOX, S. B., ICHIHARA, S., JACQUEMIER, J., LAKHANI, S. R., PALACIOS, J., RICHARDSON, A. L., SCHNITT, S. J., SCHMITT, F. C., TAN, P.-H., TSE, G. M., BADVE, S. & ELLIS, I. O. 2010. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Research* 12: 207.
- RIBALTA, T., MCCUTCHEON, I. E., ALDAPE, K. D., BRUNER, J. M. & FULLER, G. N. 2004. The mitosis-specific antibody anti-phosphohistone-H3 (PHH3) facilitates rapid reliable grading of meningiomas according to WHO 2000 criteria. *American Journal of Surgical Pathology* 28: 1532-1536.
- RICHARDSEN, E., ANDERSEN, S., AL-SAAD, S., RAKAEE, M., NORDBY, Y., PEDERSEN, M. I., NESS, N., GRINDSTAD, T., MOVIK, I., DØNNEM, T., BREMNES, R. & BUSUND, L.-T. 2017. Evaluation of the proliferation marker Ki-67 in a large prostatectomy cohort. *PLoS ONE* 12: e0186852.
- ROBERTSON, S., ACS, B., LIPPERT, M. & HARTMAN, J. 2020. Prognostic potential of automated Ki67 evaluation in breast cancer: different hot spot definitions versus true global score. *Breast Cancer Research and Treatment* 183: 161-175.
- ROBERTSON, S., RÖNNLUND, C., DE BONIFACE, J. & HARTMAN, J. 2019. Re-testing of predictive biomarkers on surgical breast cancer specimens is clinically relevant. *Breast Cancer Research and Treatment* 174: 795-805.
- ROCHE. 2023a. OptiView DAB IHC Detection Kit [Online]. Available: <https://elabdoc-prod.roche.com/eLD/api/downloads/91aa0b07-1b36-ea11-fa90-005056a772fd?countryIsoCode=au> [Zugriff 08.08.2023].
- ROCHE. 2023b. VENTANA Benchmark ULTRA [Online]. Available: <https://www.roche.de/diagnostik/produkte-loesungen/systeme/ventana-benchmark-ultra> [Zugriff 08.08.2023].
- SCHÖFFSKI, O. & EMMERT, M. 2009. Vergleichende Kostenanalyse der Nachsorge beim Mammakarzinom unter besonderer Berücksichtigung von Mistelpräparaten. *Deutsche Zeitschrift für Onkologie* 41: 119-125.
- SHUI, R., YU, B., BI, R., YANG, F. & YANG, W. 2015. An interobserver reproducibility analysis of Ki67 visual assessment in breast cancer. *PLoS One* 10: e0125131.
- SKALAND, I., JANSSEN, E. A., GUDLAUGSSON, E., KLOS, J., KJELLEVOLD, K. H., SOILAND, H. & BAAK, J. P. 2007. Phosphohistone H3 expression has much stronger prognostic value than classical prognosticators in invasive lymph node-negative breast cancer patients less than 55 years of age. *Modern Pathology* 20: 1307-1315.

- SKJERVOLD, A. H., PETTERSEN, H. S., VALLA, M., OPDAHL, S. & BOFIN, A. M. 2022. Visual and digital assessment of Ki-67 in breast cancer tissue - a comparison of methods. *Diagnostic Pathology* 17: 45.
- STALHAMMAR, G., FUENTES MARTINEZ, N., LIPPERT, M., TOBIN, N. P., MOLHOLM, I., KIS, L., ROSIN, G., RANTALAINEN, M., PEDERSEN, L., BERGH, J., GRUNKIN, M. & HARTMAN, J. 2016. Digital image analysis outperforms manual biomarker assessment in breast cancer. *Modern Pathology* 29: 318-329.
- SUN, Y., LIANG, F., CAO, W., WANG, K., HE, J., WANG, H. & WANG, Y. 2014. Prognostic value of poorly differentiated clusters in invasive breast cancer. *World journal of surgical oncology* 12: 310.
- TAGLIABUE, G., FABIANO, S., CONTIERO, P., BARIGELLETTI, G., CASTELLI, M., MAZZOLENI, G., BOSCHETTI, L., FANETTI, A. C., PUPPO, A., MUSOLINO, A., CIRILLI, C., SEGHINI, P., MANGONE, L., CALDARELLA, A., LOTTI, F., MAZZUCCO, W., BENEDETTO, A., DINARO, Y. M., SFERRAZZA, A., PINNA, P., PEROTTI, V. & AIRTUM WORKING, G. 2021. Molecular Subtypes, Metastatic Pattern and Patient Age in Breast Cancer: An Analysis of Italian Network of Cancer Registries (AIRTUM) Data. *Journal of Clinical Medicine* 10: 5873
- TAPIA, C., KUTZNER, H., MENTZEL, T., SAVIC, S., BAUMHOER, D. & GLATZ, K. 2006. Two mitosis-specific antibodies, MPM-2 and phospho-histone H3 (Ser28), allow rapid and precise determination of mitotic activity. *American Journal of Surgical Pathology* 30: 83-89.
- TETZLAFF, M. T., CURRY, J. L., IVAN, D., WANG, W. L., TORRES-CABALA, C. A., BASSETT, R. L., VALENCIA, K. M., MCLEMORE, M. S., ROSS, M. I. & PRIETO, V. G. 2013. Immunodetection of phosphohistone H3 as a surrogate of mitotic figure count and clinical outcome in cutaneous melanoma. *Modern Pathology* 26: 1153-1160.
- TRAN, B. & BEDARD, P. L. 2011. Luminal-B breast cancer and novel therapeutic targets. *Breast Cancer Research* 13: 221.
- VAN DIEST, P. J., BAAK, J. P., MATZE-COK, P., WISSE-BREKELMANS, E. C., VAN GALEN, C. M., KURVER, P. H., BELLOT, S. M., FIJNHEER, J., VAN GORP, L. H., KWEE, W. S. & ET AL. 1992. Reproducibility of mitosis counting in 2,469 breast cancer specimens: results from the Multicenter Morphometric Mammary Carcinoma Project. *Human Pathology* 23: 603-607.
- VAN STEENHOVEN, J. E. C., KUIJER, A., KORNEGOOR, R., VAN LEEUWEN, G., VAN GORP, J., VAN DALEN, T. & VAN DIEST, P. J. 2020. Assessment of tumour proliferation by use of the mitotic activity index, and Ki67 and phosphohistone H3 expression, in early-stage luminal breast cancer. *Histopathology* 77: 579-587.
- VILLANI, V., MAHADEVAN, K. K., LIGORIO, M., FERNÁNDEZ-DEL CASTILLO, C., TING, D. T., SABBATINO, F., ZHANG, I., VANGEL, M., FERRONE, S., WARSHAW, A. L., LILLEMOR, K. D., WARGO, J., DESHPANDE, V. & FERRONE, C. R. 2016. Phosphorylated Histone H3 (PHH3) Is a Superior Proliferation Marker for Prognosis of Pancreatic Neuroendocrine Tumors. *Annals of Surgical Oncology* 23: 609-617.

- WANG, Y., ACS, B., ROBERTSON, S., LIU, B., SOLORZANO, L., WAHLBY, C., HARTMAN, J. & RANTALAINEN, M. 2022. Improved breast cancer histological grading using deep learning. *Annals of Oncology* 33: 89-98.
- WOO, J. S., MOATAMED, N. A., SULLIVAN, P. S., LU, D. Y., CALLAHAN, R. & APPLE, S. 2015. Comparison of Phosphohistone H3 Immunohistochemical Staining, Ki-67, and H&E Mitotic Count in Invasive Breast Carcinoma. *Journal of Breast Cancer Survival* 1: 1.
- XI, G., HE, J., KANG, D., XU, S., GUO, W., FU, F., LIU, Y., ZHENG, L., QIU, L., LI, L., WANG, C. & CHEN, J. 2021. Nomogram model combining macro and micro tumor-associated collagen signatures obtained from multiphoton images to predict the histologic grade in breast cancer. *Biomedical Optics Express* 12: 6558-6570.
- XIANG, Z., HE, Q., HUANG, L., XIONG, B. & XIANG, Q. 2022. Breast Cancer Classification Based on Tumor Budding and Stem Cell-Related Signatures Facilitate Prognosis Evaluation. *Frontiers in Oncology* 11: 818869
- XU, A., CHU, X., ZHANG, S., ZHENG, J., SHI, D., LV, S., LI, F. & WENG, X. 2022. Development and validation of a clinicoradiomic nomogram to assess the HER2 status of patients with invasive ductal carcinoma. *BMC Cancer* 22: 872.
- YERSAL, O. & BARUTCA, S. 2014. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World Journal of Clinical Oncology* 5: 412-424.
- YU, J., WU, J., HUANG, O., HE, J., ZHU, L., CHEN, W., LI, Y., CHEN, X. & SHEN, K. 2021. A nomogram to predict the high-risk RS in HR+/HER2-breast cancer patients older than 50 years of age. *Journal of Translational Medicine* 19: 75.
- ZHU, J., LIU, M. & LI, X. 2022. Progress on deep learning in digital pathology of breast cancer: a narrative review. *Gland Surgery* 11: 751-766.

7 Erklärung zu Eigenanteil

Die Arbeit wurde im Institut für Pathologie und Neuropathologie unter Betreuung von Prof. Dr. Falko Fend durchgeführt.

Die Konzeption der Studie erfolgte in Zusammenarbeit mit Dr. Ivonne Montes und Prof. Dr. Annette Staebler.

Die immunhistochemischen Färbungen wurden durch Christiane Stoffregen und Claudia Hermann durchgeführt.

Die visuelle Auswertung der Immunhistochemie erfolgte in Zusammenarbeit mit Dr. Ivonne Montes und Prof. Dr. Annette Staebler.

Die statistische Auswertung sowie die Erstellung der Abbildungen erfolgte in Zusammenarbeit mit dem Institut für Biometrie, namentlich Dr. Lina Serna Higuita, und mit der Unterstützung von Dr. Ivonne Montes und Dr. Veronika Bahlinger.

Zur sprachlichen und stilistischen Überprüfung des Textes habe ich das KI-Tool ChatGPT als Korrekturhilfe genutzt. Ich versichere jedoch, dass inhaltliche Ausarbeitungen, Analysen und wissenschaftliche Schlussfolgerungen ausschließlich meiner eigenen Arbeit entstammen und ich keine weiteren als die von mir angegebenen Quellen verwendet habe.

Tübingen, den/Unterschrift Doktorand

8 Veröffentlichungen

Im Rahmen des GRK 2543 „Intraoperative multisensorische Gewebedifferenzierung in der Onkologie“ wurde die in dieser Arbeit erstellte „Validierungskohorte“ in einem Nachfolgeprojekt von Hannes Armbruster weiter analysiert, was zu folgender Publikation führte:

Armbruster H, Schotte T, Götting I, Overkamp M, Granai M, Volmer LL, Bahlinger V, Matovina S, Koch A, Dannehl D, Engler T, Hartkopf AD, Brucker SY, Bonzheim I, Fend F, Staebler A, Montes-Mojarro I. Aberrant p53 immunostaining patterns in breast carcinoma of no special type strongly correlate with presence and type of TP53 mutations. *Virchows Archiv* 2024 Oct 485(4): 631-642. doi: 10.1007/s00428-024-03897-3.

9 Danksagung

Mein Dank gilt allen Personen, die an der Entstehung dieser Arbeit aktiv oder passiv beteiligt waren.

Ich bedanke mich bei allen Beteiligten des Graduiertenkollegs für die gute Zusammenarbeit und die interessanten Work-Shops, denen ich im Rahmen dieser Arbeitsgruppe beiwohnen durfte.

Mein Dank gilt außerdem Prof. Dr. Falko Fend, Prof. Dr. Annette Staebler, Dr. Ivonne Montes, Dr. Veronika Bahlinger, Dr. Massimo Granai und allen weiteren Personen des pathologischen Instituts, die mich im Rahmen dieser Arbeit unterstützt haben.

Besonders bedanken möchte ich mich bei Christiane Stoffregen und Claudia Hermann, die diese Arbeit durch die Anfertigung der Immunhistochemischen Färbungen erst ermöglicht haben.

Außerdem möchte ich mich bei meiner Familie, besonders meinen Eltern und der Familie Lepschy-Stähle, für die große und bedingungslose Unterstützung während des gesamten Arbeitsprozesses herzlich bedanken.

Abschließend gilt der größte Dank Sina Stähle. In der gesamten Zeit konnte ich zu jeder Zeit auf ihren Rückhalt bauen, sowohl im Institut als auch in allen anderen Lebenslagen. Eine bessere (Labor-) Partnerin hätte ich mir nicht wünschen können.