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

Comparison of visual and digital assessment of mitotic count, Ki-67 and pHH3 immunostaining in breast cancer

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Silimon, Diana Andrea

2024

Dekan:	Professor Dr. B. Pichler
1. Berichterstatter:	Professor Dr. F. Fend
2. Berichterstatter:	Professor Dr. med. B. Hirt

Tag der Disputation: 23.04.2024

Table of contents

Illustration directory

List of tables

Abbreviations

AI	Artificial Intelligence
ASAP	Automated Slide Analysis Platform
bmp	Bitmap Image File
BRCA1/2	breast cancer gene 1/2
CI	confidence interval
CNN	Convolutional Neural Network
CogM	Cognition Master Professional Suite
DA	digital assessment
EET	extended endocrine therapy
ER	Estrogen receptors
FISH	Fluorescence In Situ Hybridization
G0	gap phase/resting phase
G1	gap phase 1
G2	gap phase 2
HER-2	human epidermal growth factor receptor 2
HPF	high power fields
ICC	intraclass correlation coefficient
IHC	immunohistochemistry
IKWG	International Ki-67 in Breast Cancer Working Group
IQR	interquartile ranges
IRS	Immune Reactive Score
jpg	Joint Photographic Experts Group
MAI	mitotic activity index
NST	No special type
PAM	Partitioning around Medoids
pHH3	Phosphohistone H3
png	Portable Network Graphics
PR	Progesterone receptor
S	synthesis phase
SCC	Spearman Correlation Coefficient

SEER	Surveillance, Epidemiology, and End Results
SI	Staining intensity
tif	Tagged Image File Format
TNBC	triple negative breast cancers
VA	visual assessment
WHO	Wold Health Organization

1. Introduction

The most widely used histologic grading system for breast cancer is the grading according to Elston and Ellis, a modification of the Bloom and Richardson grading system[1,2]. The Elston and Ellis Classification is applied on hematoxylin and eosin (HE) stained slides, the standard staining in histopathologic diagnostics, and combines details of cell morphology by the degree of nuclear pleomorphism, with a measurement of differentiation through evaluation of tubule formation and an assessment of mitosis[1]. By assigning scores to each feature, the overall score determines the histological grade.

Mitotic figures in a tissue sample may be evaluated either by mitotic count or by mitotic activity index. The mitotic count is a measure of the number of mitotic figures per 10 high-power fields (HPF), the areas showing the highest number of mitoses within the tumor, assessed through observation.

The mitotic activity index (MAI) is calculated by dividing the number of mitotic figures by the total number of cells in the tissue sample and expressing the result as a percentage.

Multiple studies have shown histological grading to be of important prognostic significance[3-5] and it recently has been incorporated globally in breast cancer guidelines[6,7]. The visual evaluation of mitotic figures is a widely and routinely used and accepted technique, as it provides a simple and rapid overview of tumor characteristics, such as growth rate and tumor behavior.

Despite all the advantages of the visual assessment of mitotic figures in histological sections, it still suffers from limited reproducibility[8-13].

The mitotic count is considered to be a subjective measure, because it depends on the interpretation of the person performing the analysis, and difficulties may arise. During the evaluation process, the pathologist must identify and count the number of mitotic figures, which can be challenging, as mitoses may appear in different forms, depending on which part of the cell cycle they are currently in.

Furthermore, they may be difficult to distinguish from other structures, such as apoptotic or karyorrhectic cells. The assessment is usually performed in a small area of the tumor, where the choice of the area is the matter of subjective

interpretation. In addition, the correct identification depends on the quality of tissue processing, differences in fixation or thickness of the sections[14-16].

To reduce interobserver variability, mitotic figure identification could be facilitated by the use phosphorylated histone H3 (pHH3) immunohistochemistry (IHC), which has been shown to be a specific marker for mitosis[17].

PHH3 is a protein component of nucleosomes in eukaryotic cells. Phosphorylation of histone H3 is a specific step during mitosis and is absent in other phases of the cell cycle. PHH3 expressions have shown promising results in multiple studies concerning various tumors[18-21].

PHH3 staining allows the identification of prophase nuclei, clear identification of mitotic figures, and leaves the apoptotic, necrotic or karyorrhectic cells unstained. Additionally, it allows faster and easier detection of mitotically active areas. Several studies have suggested that pHH3 based mitotic index may be the strongest prognostic variable in early-stage breast carcinoma[22-24].

Ki-67 is used as a marker of cell proliferation. Ki-67 is a protein that is expressed at high levels during all phases of the cell cycle, except for the resting phase G0 and is often used as an indicator of cell growth and division. Ki-67 immunohistochemistry is scored as a proliferation index, based on the percentage of positive cells in a tissue sample. The proliferation marker has been shown to be a good objective substitute for mitotic counts when used in a grading system[25] and has also been valued as a prognostic factor being associated with breast cancer outcomes[26,27]. Ki-67 previously has been already considered as a biomarker for therapeutic decisions[28]. The reproducibility of proliferation indexes of $\leqslant$10% and $\geqslant$35%, has been shown to be much better than in the intermediates, but since there is no general cut-off value for Ki-67[29-31], the clinical significance of its expression may vary, depending on the specific context of other clinical factors and patient cohort of a study.

PHH3 IHC and the Ki-67 proliferative index are quantifiable markers and therefore amenable to computer-assisted image analysis.

In recent years, technological innovations have offered new, reliable, and accurate approaches for a more objective assessment of tumor tissue.

With the use of whole slide imaging (WSI), which means scanning of slides at a high resolution, it has been made possible to view and analyze each slide on the computer. Image analysis software may be used to perform quantitative image analysis on digital pathology images, or to measure various features of cells and tissues and to classify them based on their characteristics. Improvements in diagnostic reproducibility have previously been achieved through computer-assisted image analysis[32].

Artificial intelligence (AI) and machine learning algorithms are increasingly being trained to recognize patterns and features that are indicative of breast cancer and to classify tissue samples. There is ongoing research to develop AI based diagnostic tools to assist pathologists in the analysis of histological images[33].

In this study the quantitative image analysis software Cognition Master Professional Suite (CogM) was used. It is a collection of image analysis software tools for virtual microscopy, evaluation, and analysis of scanned digital histological slides. One of the CogM modules, the Ki-67 Quantifier, has been previously validated in a neoadjuvant breast cancer clinical trial as a computer-based approach for the quantitative evaluation of Ki-67 scoring based on a cell detection method[34].

The main goal of our study was to compare probable changes in tumor grading, agreement, and correlation between visual and digital methods for assessing mitotic count, and the quantifiable marker pHH3 and Ki-67 in breast carcinoma tissue and thus reducing the subjectivity and variability of histological grading.

1.1. Epidemiology

Breast cancer is the most prevalent cancer in women globally, with 2.3 million cases diagnosed in 2020, and remains the leading cause of cancer-related deaths, accounting for 685.000 deaths worldwide[35].

The lifetime risk of developing breast cancer in women is approximately 1 in 8, and the incidence of breast cancer increases with age. Risk factors for breast cancer include genetic mutations such as BRCA1/2, a family history of breast or ovarian cancer, alcohol consumption, and obesity after menopause.

Despite the high incidence of breast cancer, the death rate has been declining since the 1990s, due to screening, prevention, early detection, more awareness and continuous improvement in treatment[36]. Early detection through preventative examinations and mammography has led to the detection of breast cancer in earlier stages, smaller in size and as lymph node negative tumors at diagnosis, thereby increasing the importance of histological grade in breast cancer diagnosis and treatment[5,37].

1.2. Classification

Breast carcinoma represents a heterogeneous group of diseases characterized by distinct clinical, histopathological, and molecular characteristics. Morphological subtypes have traditionally been used to classify invasive breast cancer, which includes various biologically distinct entities, each with unique pathological features and clinical presentations.

Patient age, along with tumor size, lymph node involvement, metastasis, histological type and grade are clinical-pathological parameters that are crucial for breast cancer assessment. Heterogeneity in histopathological and biological characteristics in breast carcinomas results in varying responses to treatment. [38,39].

Immunohistochemical markers are often used to guide treatment decisions, making individual treatment possible. To further refine breast cancer classification, gene expression profiling has been used to develop molecular subgroups[40].

To aid the restricted daily clinical application of gene expression profiling, due to expenses and the time-consuming nature of the technique, immunohistochemistry is used as a time efficient surrogate[41,42].

The molecular subgroups, in combination with conventional prognostic indicators such as TNM staging, tumor size, and histological grading, represent relevant prognostic and predictive factors that have significant implications for treatment and affect therapeutic strategies[7].

The following chapters are focusing on giving an overview of the morphological classification and histological grade.

1.3. Morphological classification

Breast cancer comprises a broad spectrum of types and subtypes, each of which is associated with varying prognoses and treatments. According to the WHO classification of 2012, the majority of breast cancer cases are carcinomas[43], which arise from epithelial cell-based components of the inner layer of the terminal ducts and lobules. Sarcomas, which arise from connective tissues such as blood vessels or myofibroblasts, represent approximately 1% of breast cancer cases and comprise the second category. Carcinomas are further classified into in situ and invasive carcinomas. In situ carcinomas are characterized by the presence of abnormal cells confined within the mammary ducts and without invasion of the basement membrane. In contrast, invasive carcinomas have breached the ductal walls and invaded the surrounding tissues. If left untreated, invasive carcinomas have the potential to metastasize to other tissues and organs. The majority of invasive breast cancers, accounting for approximately 80%, lack distinct histological features and are categorized as invasive carcinoma of no special type (NST). This group was previously known as invasive ductal carcinoma until the update in the WHO classification in 2020[43].

1.4. The histological grade

The Elston and Ellis Score, also known as the Nottingham grading system, has become widely adopted as a global standard for evaluating breast carcinoma and guiding treatment decision. Histologic grade is a measure of the degree of differentiation based on the evaluation of three morphological features: tubule formation, nuclear pleomorphism, and mitotic nuclei. A study by Schwartz et al. analyzed a large cohort of breast cancer cases from the Surveillance, Epidemiology, and End Results (SEER) program and demonstrated that histological grading remains a significant prognostic factor for overall survival[44]. This finding highlights the continued clinical relevance and importance of incorporating histological grade into the disease management of breast cancer.

Table 1: Elston and Ellis grading system

Tubule formation	
1 point	Tubule formation in >75% of the tumor
2 points	Tubule formation in 10% to 75% of the tumor
3 points	Tubule formation in <10% of the tumor
Nuclear pleomorphism	
1 point	Small, regular, uniform cells, nucleus shapes only differ mildly
2 points	Moderate nuclear size and variations
3 points	Nucleus shape and size difference significantly
Mitotic count/field area (0.159mm^2)	
1 point	0-5
2 points	6-11
3 points	>12
Final score	
3,4,5	Grade I: well differentiated
6,7	Grade II: moderately differentiated
8,9	Grade III: poorly differentiated

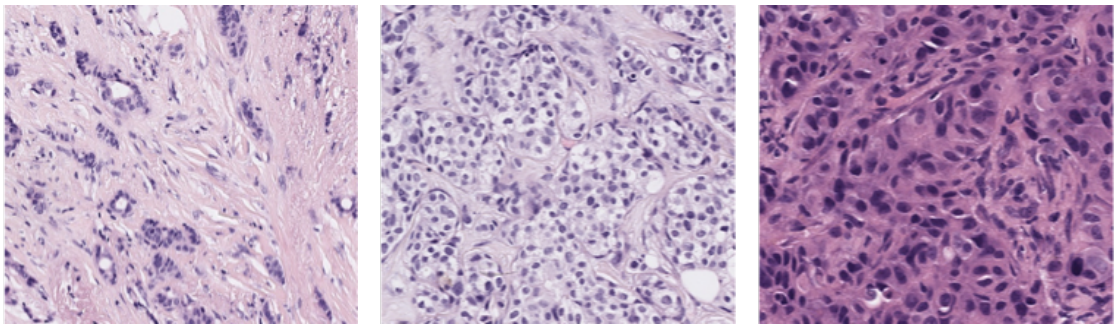
The Elston and Ellis Classification is conducted on HE-stained slides and is evaluated by a numeric scoring system of 1-3 in each category. The grades are classified based on the total scores.

The cut-off values for mitotic counts are determined by the diameter of the high-power field and its corresponding area[7].

In this study, the mitotic count was scored as the total number of mitotic figures in non-overlapping 10 consecutive HPFs with a field diameter of 0.45 mm and the area assessed is 0.159 mm^2 (Table 1).

Counting is preferred in tumor rich areas. One of the challenges of mitosis detection on hematoxylin and eosin-stained slides is distinguishing true mitotic figures from apoptotic bodies, tissue artifacts, dark staining or dark nuclei[12,13].

The total score is obtained by adding the scores on tubule formation, nuclear pleomorphism and mitotic index. The histological grades are determined by the total scores.



a) b) c)

Figure 1: Breast cancer of no special type is graded according to the Elston and Ellis Score, which is based on the degree of three morphological features: tubule formation, nuclear pleomorphism, and mitotic count (Table 1). Figure 1 displays examples of histological samples of the different grades: a) grade I: well differentiated, b) grade II: moderately differentiated and c) grade III: poorly differentiated breast tissue.

1.5. Immunohistochemistry

The molecular classification of breast cancer began with the identification of hormone receptor markers. Estrogen receptors (ER) were first discovered in the 1960s by Jensen and Jacobson[27] and later studies in the 70s by McGuire et al. demonstrated that the ER was detectable in 60-80% of breast cancers[45]. In the 1980s, amplification of HER-2 (Human Epidermal Growth Factor 2) was shown in 30% human breast cancer[46]. Subsequently, gene expression profiling was developed, which allowed for molecular classification of breast cancer into subgroups[40]. The luminal class was further divided into luminal A and luminal B, resulting in the identification of an additional fifth class of breast cancer: luminal A, luminal B HER2-positive and luminal B negative, HER2-enriched and triple negative breast cancer[47]. In clinical practice, immunohistochemical analyzes of tumors based on status of ER, PR, and HER2 are used to determine molecular

subtypes. This method is easier and more cost-effective than gene expression profiling and provides similar results[48].

1.5.1. Hormone receptor biomarkers

In the present day, the utilization of immunohistochemistry for the evaluation of breast cancer has become a common practice for the analysis of steroid receptor status[49]. The majority of human breast cancers, approximately 70-80%, are hormone-dependent and estrogen receptor positive[50,51] and IHC testing can effectively detect estrogen and progesterone receptors (ER/PR) in cancer cells from tissue samples. These receptors are dependent on estrogen or progesterone for their growth, and patients with ER and PR positive breast cancer usually respond well to endocrine therapy[51].

1.5.2. Human Epidermal Growth Factor 2

Human Epidermal Growth Factor 2 (HER2) is a gene which encodes a transmembrane tyrosine kinase and is a crucial regulator of cellular growth and proliferation in normal breast tissue. However, in about 15-20% of breast cancer cases, HER2 gene amplification or overexpression occurs[52]. The immunohistochemistry has been adopted as the screening test and is a routine practice for invasive breast cancer in pathology. The IHC test involves staining the tumor tissue sample with antibodies that specifically target the HER2 protein. The tissue is then examined under a microscope and given a score. Scores of 0 or 1+ indicate HER2-negative tissue, while a score of 3+ indicates HER2-positive tissue. A strong, circumferential membrane staining is required for a positive (3+) result. However, if the tissue is scored as 2+, it is considered equivocal, and further testing using fluorescence in situ hybridization (FISH) is recommended. Patients with HER2-positive breast cancer are typically treated with targeted therapies, such as trastuzumab (Herceptin), which specifically target the HER2 protein[52].

1.5.3. Proliferation index

The proliferation index describes what proportion of the cancer cells within the tumor are growing and increasing in number. A higher number indicates a more aggressive tumor[53]. Immunohistochemical analysis of the Ki-67 index represents the cell proliferation of the cancer. Ki-67 is a nuclear protein which was found by Gerdes et al. in 1983 in a Hodgkin lymphoma cell line[54]. The most widely used proliferation marker in breast cancer is Ki-67[31]. Ki-67 protein is present during all phases of the cell cycle (G1, S; G2 and mitosis) but is absent during the resting phase, G0[55]. Ki-67 level varies during the cell cycle. Levels are low in G1 and early S phase and reach the maximum in early mitosis. Rapid decrease in Ki-67 rates occur in anaphase and telophase. MIB-1 is a monoclonal antibody that reacts with the Ki-67 nuclear antigen and can be used on both frozen and paraffin-embedded section[56]. Ki-67 may be used to distinguish between low- and high-proliferating tumors and its reproducibility has been shown to be more consistent when the rate of cell division is either very low ($\leqslant$10%) or very high ($\geqslant$35%), but there is still no general cut-off value for Ki-67[29-31].

Furthermore, the proliferation index may be used to guide clinical decisions regarding adjuvant chemotherapy in ER-positive tumors and is predictive of responsiveness to neoadjuvant chemotherapy[57,58].

Currently monitoring Ki-67 levels during or after neoadjuvant endocrine or chemotherapy helps to make decisions on the effectiveness of the current treatment or if the treatment plan needs to be adjusted. A high Ki-67 index is predictive of responsiveness to neoadjuvant chemotherapy and for potential effectiveness of the treatment and a decline of Ki-67 expression after neoadjuvant endocrine therapy predicts better disease-free survival[59].

Moreover, patients with rapidly proliferating breast cancer benefit from adjuvant chemotherapy as opposed to those with slowly proliferating tumors[26].

Furthermore, Ki-67 may have a role as part of the IHC4 score, based on quantitative assessment of four immunohistochemistry markers (ER, PR, HER2, and Ki67[60]). The IHC4 score can be applied to postmenopausal women with early-stage breast cancer who have been treated with 5 years of hormonal therapy to estimate the risk of recurrence at 10 years and the potential benefits

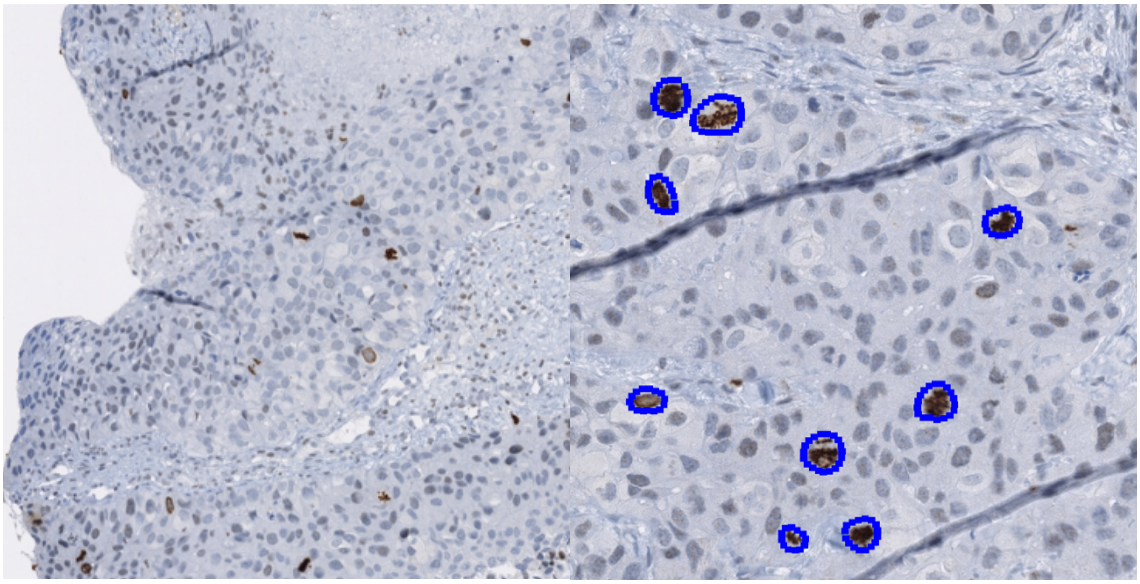
of extended endocrine therapy (EET) beyond the standard 5 years of treatment with hormonal therapy[61]. The score can help guide clinical decision-making by identifying patients who have a low risk of recurrence and would not benefit from extended therapy[61].

1.5.4. Phosphohistone H3

Histone H3 is a nuclear core histone protein of DNA chromatin and plays an important role in chromatin condensation during mitosis after phosphorylation on Serin 10[62]. Phosphorylation appears to be essential for maintaining the compact chromosome state and is involved in controlling DNA replication during the cell cycle. It occurs during late G2 (second gap phase, where the cell undergoes growth and prepares for cell division) and early prophase and dephosphorylation from late anaphase to early telophase. In metaphase histone H3 is always heavily phosphorylated and resulting in a high level of pHH3 expression. During interphase there is minimal to no pHH3 expression[58]. PHH3 allows for better distinctions between mitotic cells and mitotic mimickers such as apoptotic or karyorrhectic cells, as it only stains actively dividing cells[63]. The assessment of pHH3 staining is uncomplicated and straightforward, given the contrast between brown-stained mitotic cells and blue-stained non-mitotic cells aiding in recognition.

Multiple studies have verified pHH3 concerning various tumors, such as thin melanoma, meningiomas, pancreatic neuroendocrine carcinoma, pulmonary neuroendocrine carcinoma, and astrocytomas for its sensitive and specific role as a marker of mitotic figures [18-21,64].

PHH3 IHC has been shown as a dependable and sensitive technique for identifying mitotic figures and hot spots for mitotic activity assessment[63,65]. In a study conducted by Cui et al., they observed a significant correlation between the mitotic indices gathered from pHH3 staining and those obtained from HE staining and the authors suggested the possibility of integrating the pHH3 marker into breast cancer grading[66].



a) b)

Figure 2: Examples of mitotic figures with pHH3 staining: the slide appears in different shades of blue with dark-brown-stained cells, indicating mitotic indices expressing pHH3.
a) Overview of a pHH3-stained slide without annotations.
b) Close-up of a pHH3-stained slide with blue encircled mitotic indices.

1.6. Molecular classification

Breast cancer is classified based on classical immunohistochemistry, genetic, and molecular characteristics[40]. Subtypes of breast cancer have been identified using advanced techniques such as gene expression profiling with complementary DNA microarrays[67]. In clinical practice the status of immunohistochemistry staining such as ER, PR and HER2 are used to determine these molecular subtypes. The molecular classification of breast cancer provides more precise information for personalized therapeutic decision-making and prognosis[68].

The St. Gallen Consensus of 2011 has classified breast cancer into four subtypes based on molecular characteristics, namely Luminal A (ER+/PR+/HER2/low Ki-67), Luminal B (ER+/PR+/HER2-/+/high Ki-67), HER2-overexpression (ER-/PR-/HER2+) and triple negative breast cancers/TNBCs (ER-/PR-/HER2-)[69].

Table 2: An overview of the molecular subtypes according to IHC staining

Subtypes	ER	PR	HER2	Ki-67
Luminal A	ER positive	PR positive/negative	Negative	Low
Luminal B	ER positive	PR positive/negative	Negative	High
Luminal B	ER positive	PR positive/negative	Positive	Any
HER2-enriched	Negative	Negative	Positive	Any
Triple negative	Negative	Negative	Negative	Any

1.6.1. Luminal A

Luminal A represents the most prevalent molecular subtype of breast cancer, accounting for 30-45% of all newly diagnosed cases. According to the St. Gallen International Expert Consensus of 2013, it is characterized by an estrogen receptor (ER) expression of more than 1%, progesterone receptor (PR) expression of more than 20%, human epidermal growth factor receptor 2 (HER2) expression of =/<10%, and low to medium Ki-67 expression. Luminal A subtype is associated with a favorable prognosis compared to the other subtypes[70]. Given the positive hormone receptor status, recommended treatment for Luminal A breast cancer includes endocrine therapy such as selective estrogen receptor modulators (e.g., tamoxifen) or aromatase inhibitors (e.g., anastrozole)[71].

1.6.2. Luminal B

The subtype Luminal B comprises approximately 20% of all cases and is further divided into Luminal B HER2 positive and negative, depending on the status of ER, PR and HER2 receptors, with high Ki-67 expression[71]. Compared to Luminal A, Luminal B breast cancer is slightly more aggressive and has intermediate prognosis. The recommended treatment differs for Luminal B HER2 positive and negative cases. For Luminal B HER2 negative cases, the recommended treatment is endocrine therapy, with the option of additional systemic therapy. In Luminal B HER2 positive cases, anti-HER2 therapy (such as trastuzumab) is also needed in addition to endocrine therapy and chemotherapy[71].

1.6.3. HER2 positive

Approximately 15-20% of all breast cancers are classified as HER2-positive. This subtype is characterized by HER2 positivity, lack of steroid receptors (ER<1% and PR<20%) but high expression of Ki-67. Amplification of the HER2 gene is associated with faster growth, higher mitotic index, higher histological grade, and worse prognosis[72]. The recommended treatment for HER2-positive breast cancer is HER2 blockers, either with humanized monoclonal antibodies (trastuzumab) or molecular receptor tyrosine kinase inhibitors (lapatinib).

1.6.4. Triple negative

Triple negative breast cancer (TNBC) accounts for approximately 10-20% of all breast cancer cases and is characterized by the absence of steroid receptors (ER<1% and PR<20%) and HER2 proteins, but with a high proliferative index. TNBC is generally considered more aggressive than other types of breast cancer, with a higher risk of metastasis and recurrence. This subtype is more common in women with BRCA1 gene mutations, among women younger than 40 years and African American women. The standard therapy for TNBC without distant metastases is currently anthracycline-/taxane-based chemotherapy, preferably administered neoadjuvantly[73,74].

1.7. Digital pathology

At present, the gold standard for evaluating histological slides in pathology laboratories is through visual examination by a pathologist. This method involves the evaluation of histological images under a microscope. However, the current method of assessing histological slides by visual examination is subjective and prone to inter-observer variability, resulting in diagnostic discrepancies. To overcome this issue, computerized image analysis methods have been proposed as an objective and quantifiable solution with the potential to improve reproducibility.

Whole slide imaging is a commonly used method by pathologists worldwide, where conventional glass slides are scanned to produce digital slides. These scanners allow for a magnification of the slides at 20x or 40x with 0.46 µm/pixel or 0.23 µm/pixel spatial resolution. Spatial resolution is indicative of the amount of detail the scanner can capture, with higher resolution scanners providing greater pixel definition. WSI scanners by manufacturers such as Ventana, Aperito, Hamamatsu, and Leica are among the most commonly used in pathology laboratories. The digitization of histopathological images has enabled the automation of breast cancer classification using computer-assisted image analysis software.

Image analysis software offer several advantages including the potential to improve accuracy and precision in the analysis of tissue samples. By providing more objective measurements, these software tools may help to reduce subjectivity and variability associated with traditional visual evaluation by a pathologist. Additionally, image analysis software has the potential to reduce the time and effort required for the analysis.

The search for better, more objective methods has led to the development of deep learning algorithms. These algorithms are capable of extracting thousands of features, including small areas, from WSI. These features, commonly referred to as tiles, are utilized to classify tissue samples into diverse categories, such as cancerous or non-cancerous. The outcomes achieved from analyzing individual tiles within a WSI are eventually integrated to generate a diagnosis for the entire slide.

Convolutional Neural Networks (CNNs) have become the preferred approach for analyzing pathology images since their introduction. These networks can analyze images and recognize features such as contours by detecting changes in the color intensity of pixels in surrounding areas. The network can identify larger and more complex patterns by combining these features across multiple layers[75-78].

In a recent study conducted by Chan et al., the current state of Artificial Intelligence development was reviewed. The authors specifically highlighted a study conducted by Barsha et al., which proposed a combination of multiple common deep learning algorithms to detect and invasive breast cancer[79,80]. The

accurate detection of mitotic count is a crucial component of histological grading, and numerous studies have been conducted in an attempt to automate this process using AI[81-83].

Despite the successes observed thus far, the practical implementation of AI may be complicated by the algorithms tendency to identify excessive information on histological slides (e.g., artifacts, ink stains). Additionally, various factors such as section thickness, staining characteristics, and imaging techniques may also influence the accuracy of the analysis[84].

1.8. Aims of this study

Aim 1: Assessing the accuracy and interobserver variability of the Elston and Ellis Score

Aim 2: Using Ki-67 and pHH3 as quantifiable immunohistochemical markers for grading assessment, reducing the subjectivity and variability of histological grading

Aim 3: Integrating digital pathology to improve the accuracy of quantitative immunohistochemical staining

2. Material and Methods

2.1. Patient cohort and samples

Archival formalin-fixed paraffin-embedded tissue blocks were obtained from the archives of the Institute for Pathology and Neuropathology in Tübingen spanning the period from 2017 to 2019. A total of 74 cases of invasive breast carcinoma of no special type (NST) were selected for inclusion, with 1-2 slides each, resulting in a dataset of 84 HE-stained slides. Each case contributed 1-2 slides, resulting in a dataset of 84 slides stained with HE. This study included multiple slides of the same tumor for some cases, when available, to enhance the range of histologic features and tumor heterogeneity observed in real clinical cases. This approach may also provide insights into the degree of variation between samples from the same individual.

For each of the 84 hematoxylin and eosin-stained slides, one corresponding immunohistochemically stained slide was used to evaluate the expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki-67. All tissue blocks had previously been stained, except for phosphorylated histone H3 (pHH3) IHC, which was specifically performed for this study.

The project was evaluated by the Ethics Committee (Ethics Committee at the Medical Faculty of the Eberhard Karls University and at the University Hospital Tübingen) and approved on 17.08.2021, the project number is 547/2021BO2.

2.2. Material

All cases were subjected to re-evaluation under a microscope (Zeiss Axio Imager A1). Additionally, digital images of the tissue slides were generated using the Ventana DP 200 slide scanner (Roche Diagnostics), which allowed for high-resolution digitization of the pathological images and immunohistochemical stains at a magnification level of 20x.

Table 3: List of devices

Name	Article, Manufacturer
Microscope	Zeiss Axio Imager A1
Slide scanner	Ventana DP 200 slide scanner, Roche Group
BenchMark Ultra	Ventana Medical Systems, Roche Group

Table 4: List of software

Software	Purpose of use	Manufacturer
Ventana Image Viewer	Virtual slide viewing	Roche Diagnostics
ASAP Automated Slide Analysis Platform	Annotation	Computation Pathology Group at Radboud University Medical Center
Cognition Master Professional Suite	Digital assessment and quantification	VMscope GmbH (Berlin)

2.3. Methods

2.3.1. Hematoxilin and eosin staining

HE-staining is a critical process for highlighting histological features. In breast cancer tissue, nuclei are stained in a deep purple or blue color and the cytoplasm has various degrees of pink staining.

The formalin fixed tissue specimens must dry for 12 minutes. Before staining, the slides were dewaxed in Xylol twice for 2,5 minutes each, then passed through descending grades of alcohol (100%, 96%, 70%) and rinsed in distilled water for 30 seconds. The sections were stained in Mayer's hemalum solution for 2 minutes first and two more times for 4 minutes. Afterwards, sections were rinsed for 10 minutes in tap water, before counterstaining them in 1% eosin solution for 1 minute. After a 10 second rinse with tap water, the sections were dehydrated with ascending grades of alcohol for 20 seconds with 70% alcohol, for 1 minute

with 96% alcohol and twice for 1 minute with 100% alcohol. Finally, Xylene was used as a clearing agent for 4 minutes.

2.3.2. Immunohistochemistry staining

The investigation of the tissue sample by immunohistochemistry provides information through specific antibodies, which are used to bind a target protein.
The bound antibody is then visualized by dye created by an enzymatic reaction.
Immunohistochemical staining was performed on formalin fixed, paraffin embedded tissue sections. The immunohistochemical staining was performed according to the BenchMark ULTRA XT (Ventana Medical Systems, Roche Group, Tucson, AZ, USA) manufacturer protocols.
The BenchMark ULTRA XT is a fully automated immunohistochemistry slide staining system and was used for IHC staining with the Ventana OptiView DAB IHC Detection Kit. The automated staining platform aids in securing and improving consistency of the IHC staining procedure.
The BenchMark Ultra XT was used to conduct all involved steps, beginning with deparaffinization, rehydration and antigen retrieval. The latter involves pretreatment of tissue to retrieve antigens masked by fixation and to make them more accessible to antibody binding. Antigen retrieval was performed with CC1 (EDTA buffer) (prediluted; pH 8.0), on the BenchMark Ultra automated slide stainer for 64 minutes at 100°C.
After antigen retrieval, the sections were incubated in 3% H_2O_2, then with the primary antibody for a fixed set of time. The primary antibodies are mouse or rabbit antibodies that bind to the target proteins. A secondary antibody is used to detect the primary antibody and only binds mouse or rabbit antibodies and does not bind human antibodies. The secondary antibody utilized is known as the OptiView HQ Universal Linker, which consists of a blend of HQ-labeled antibodies. These antibodies include goat anti-mouse IgG, goat anti-mouse IgM, and goat anti-rabbit antibodies, with HQ representing a proprietary hapten that is covalently attached to the goat antibodies
The secondary antibody was detected through an indirect method utilizing an enzyme bound tertiary antibody, a multimer (OptiView HRP Multimer) a multi

molecular structure that contains several enzymes. The enzyme used to catalyze the reactions that produce dye is horseradish peroxidase (HRP).

The mouse monoclonal anti-HQ-labeled HRP tertiary antibody was visualized by H_2O_2, DAB chromogen (3, 3'-diaminobenzidine tetrahydrochloride) and copper enhancer. DAB, when oxidized by HRP, results in the brown colored visualization of the antigen bound antibodies.

The employed primary antibodies are manufactured by Ventana (Tucson, USA) and are listed in the following table (Table 5).

Table 5: Characteristics of the antibodies used in the study

Antigen	Antibody	Clone	Dilution	Incubation time	Incubation temperature
ER	ANTI-ER (Rabbit-IgG)	SP1	Ready to use[1]	32 min	37°C
PR	ANTI-PR (Rabbit-IgG)	1E2	Ready to use	32 min	37°C
HER2	ANTI-HER-2/NEU (Rabbit-IgG)	4B5	Ready to use	32 min	Room temperature
Ki-67	ANTI-Ki-67 (Rabbit-IgG)	MIB1	1:400	32 min	37°C
pHH3	Anti-Histone H3	BC37	1:500	32 min	37°C

[1] Prediluted for diagnostic purposes

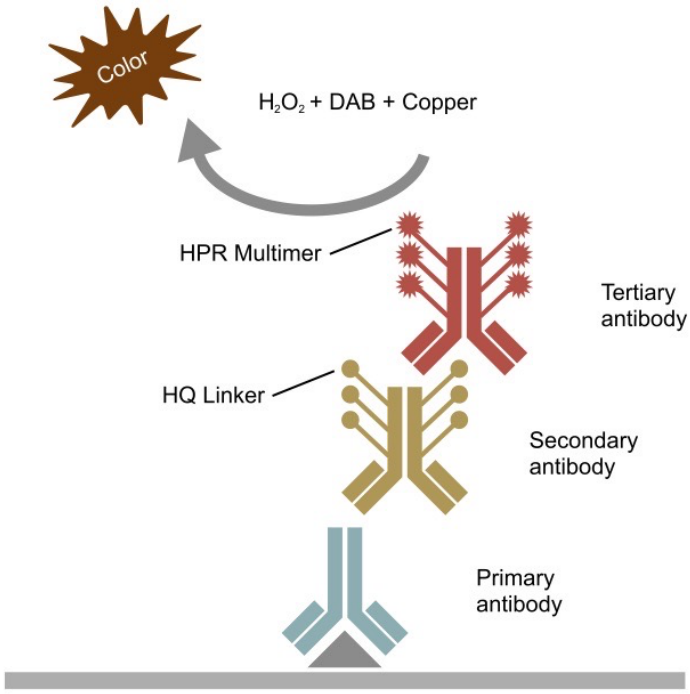


Figure 3: Image of the principle of immunohistochemistry: The primary antibody binds the antigen, and the secondary antibody is used to detect the primary antibody. The secondary antibody is detected through an indirect method utilizing an enzyme bound tertiary antibody. The enzyme-catalyzed reaction results in dye.

2.3.3. Whole slide scanning

All glass slides with tissue sections with HE- and IHC-staining were digitized using the Ventana DP 200 slide scanner (Roche Diagnostics) at 20x with a scan resolution of 0.46 µm/pixel. The Ventana slide scanner is a 6-tray based scanning system. Digital imaging vendors may use different technologies to acquire the digital images, but most use either a tile-based or a line-based system. A tile-based scanner creates hundreds of individual photos that cover the whole tissue area and are subsequently stitched together to create one seamless image. The Ventana slide scanner is a line-based scanner, creating line scans of tissue areas[85], thus creating the entire slide with fewer images and increasing the digitization speed[86].

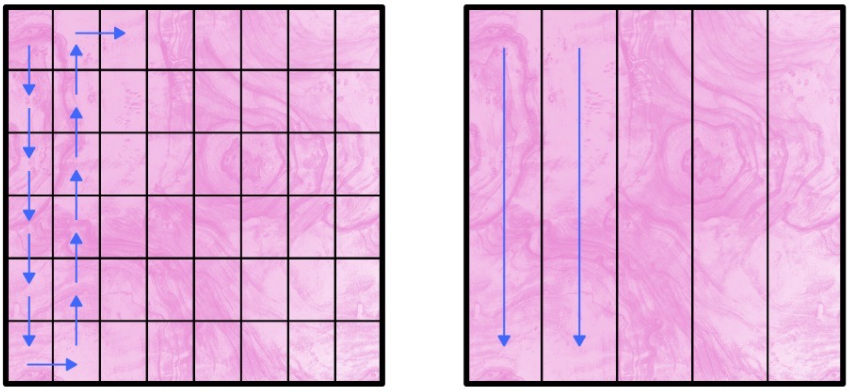


Figure 4*:* Different technologies for the acquisition of digital images**:**
a) Tile based scanning
b) Line based scanning

2.3.4. Slide management

The digital slides were examined using the Ventana Image Viewer software (Roche Tissue Diagnostics, Tucson, USA). This software facilitated efficient slide management, enabled the display of digital slides, and allowed for the incorporation of annotations as needed.

Despite the large file size of digital slides, pathologists can utilize virtual slide viewing software on their computers, because the slide viewers function by accessing and displaying only the current area of interest, eliminating the need to download the entire slide. As new areas are observed, previously requested areas are closed[87].

2.4. Pathological examination

In this study, two different methods were employed to assess the histopathological features of breast cancer tissue, which were subsequently compared for their accuracy and reliability.

The first method involved visual assessment, the traditional approach of slide evaluation of mitotic indices or positive cells by an observer under a microscope. The second method involved the use of automated digital analysis by a computer-based software. In the following, both methods are described.

2.4.1. Visual assessment

Data was collected from the pathology reports for each of the 74 cases of breast cancer included in the study. This included information on the Elston and Ellis Scores, as well as the scores for each of the three components - nuclear pleomorphism, tubular formation, and mitotic count - and IHC scores for HE, ER, PR, HER2, and Ki-67. The pathology reports were conducted by experienced pathologists at the Institute of Pathology in Tübingen.

Corresponding to the reported data, pathological re-evaluation was conducted on all samples HE, ER, PR, HER2, Ki-67 and pHH3.

The grading of the HE slides was performed according to current guidelines using the Elston and Ellis Score, which emphasizes the three main features nuclear pleomorphism, tubular formation, and mitotic count[7].

Mitoses were counted as part of the re-evaluation on glass slides under a Zeiss light microscope. Counting was performed at 400x magnification with 10x ocular and 40x objective. Mitotic count was scored as the total number of mitotic figures in non-overlapping 10 consecutive HPFs with a field diameter of 0.45 mm and the area assessed is 0.159 mm^2. Areas with ductal carcinoma in situ were excluded. PHH3 IHC was scored as the total number of positive cells in 10 consecutive HPFs where the immunolabeling was most prevalent.

The assessment of Ki-67 positivity was based on the percentage of positively stained nuclei, with a threshold of <20% used to determine positivity. ER and PR positivity were evaluated using the Immune Reactive Score (IRS) method developed by Remmele and Stegner, which assigns a total score of 2 to 12 based on the intensity and percentage of positive staining[88].

HER2 expression was assessed according to the following criteria: 0 for the cells stained in less than 10%. 1+ as >10% of the tumor cells stained positive but no staining of the cell membrane. 2+ as >10% of the tumor cells stained positively but only moderate staining of the cell membrane occurred and 3+ as >10% of the tumor cells stained positively with strong staining of the circumferential membrane.

Table 6: IRS Staining intensity (SI) evaluation method by Remmele and Stegner, assessing positive ER and PR expression. The IRS score is the result of the score for percentage of positive cells multiplied by the score for staining intensity.

Percentage of positive cells	Score	Staining intensity	Score
No positive cells	0	No staining visible	0
less than 10	1	Low staining intensity	1
10-50	2	Moderate staining intensity	2
51-80	3	High staining intensity	3
more than 80	4		
IRS Immunoreactive Score		**Score**	
Negative		0 - 1	
Positive		2 - 12	

2.4.2. Computer assisted digital assessment

For the digital image analysis, the software Cognition Master Professional Suite (VMscope, Berlin, Germany) was used to analyze the digital images based on structures, intensity, color, and texture of pixel regions. Each stained section was digitized with the Ventana slide scanner at 20x magnification. The software requires .jpg, .png, .tif or .bmp images, which were generated manually by selecting representative fields of the digital image using either the screenshot tool in Ventana picture viewer in 40x or the image generating feature in Cognition Master Professional Suite. Also, manual selection helped identifying the regions of interest, since quantitative image analysis algorithms cannot always automatically find the focused tissue regions.

2.4.3. Ki-67 Quantifier

Ki-67 scoring was conducted by the Ki-67 Quantifier module of the Cognition Master Professional Suite (VMscope). The Ki-67 Quantifier is a feature for computer assisted scoring of the Ki-67 index. To utilize the quantification tool, previously selected .png (Portable Network Graphic) images were used. By accessing the folder icon in the toolbar, the image may be uploaded through selecting the required file. Analyzing may begin with the start button, and after a short loading time, the results (the number of positive cells, the total number cells, and the percentage of positive cells) are presented. The percentage of positive cells was used to describe the Ki-67 proliferation index.

Digital assessment of pHH3 was performed using the Ki-67 quantifying tool as well, which automatically counted the total number of positive cells in each histological image (10 HPFs).

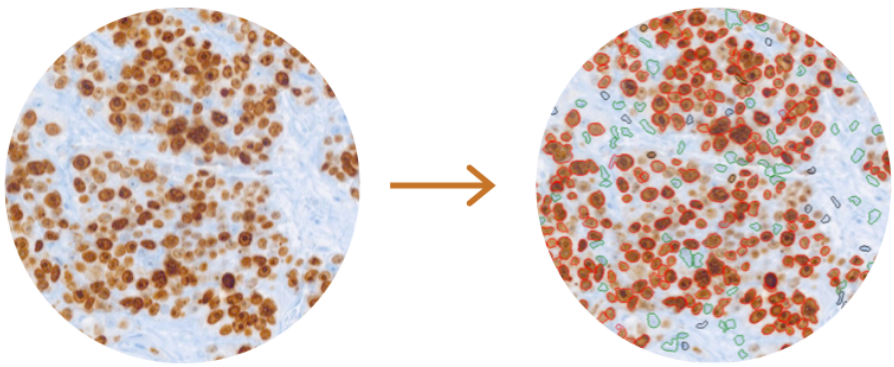


Figure 5: a) Tissue sections stained for Ki67 with MIB1 antibody (brown stain) and counterstained with Mayer's hematoxylin (blue stain)
b) Ki-67-stained tissue sample after analysis with the Ki-67 Quantifier (the cells circled in green: Ki67-negative tumor cells; the cells circled in red: Ki67-positive tumor cells; the cells circled in black: non-tumor cells)

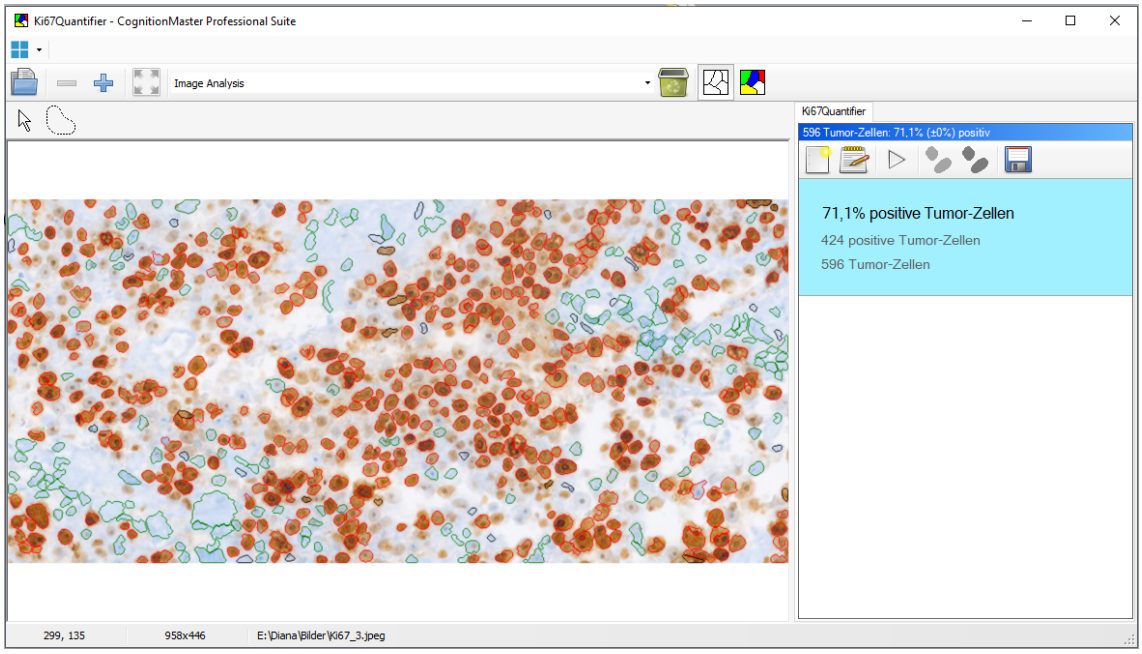


Figure 6: Ki-67-stained slide uploaded and opened the Ki-67 Quantifier module of CogM

2.4.4. ER and PR Quantifier

Cognition Master Professional Suite also offers an ER and PR quantifying tool. The automated scoring tool provides a score either according to Remmele or to Allred. For this database (ER=84 slides and PR=84 slides) the scoring according to Remmele was used. Therefore, not only the amount of stained tumor cells were counted, but also the intensity of the staining was assessed for the final score (Table 6).

2.5. Annotations

The annotation of images was performed using the Automated Slide Analysis Platform (ASAP) (Computation Pathology Group at Radboud), which allows the user to draw directly onto whole slide images and export them for further use. In total, 84 HE-stained slides and 84 pHH3-stained slides were digitized and annotated. Annotations were made on the digitized images using the tools provided by the ASAP software. All cancerous tissue was annotated. Tissues with NST invasive carcinoma show a distortion of the architecture as well as higher

nuclei variability and an increasing number of mitoses. In contrast, normal tissue has maintained architecture and well-differentiated nuclei. The mitotic count was assessed and scored according to the Elston and Ellis Classification on HE-stained slides. On pHH3-stained slides, positive cells were counted in 10 high power fields.

2.5.1. Development of a training dataset for a deep learning algorithm

For the purpose of developing a cancer detection algorithm, cancerous regions and mitoses were annotated in order to create an effective training dataset. The dataset used for this study was prepared as a training set for a deep learning algorithm developed by Simon Holdenried-Krafft and Prof. Dr. Ing. Lensch from the Department of Computer Science in Tübingen as part of the project GRK 2543: Intraoperative Multisensory Tissue Differentiation in Oncology.
The deep learning algorithm was trained on this dataset to learn the automated distinction between cancerous and non-cancerous tissue, therefore cancerous tissue was annotated on the HE slides with NST in ASAP. The annotations were controlled by two experienced pathologists. Areas with ductal in situ carcinoma, tissue artifacts, necrosis were excluded. To simulate the mitosis counting as defined in the Elston and Ellis Grading system, 10 regions of mitosis were determined per tissue sample. In these regions, all mitosis were encircled and counted. Positive cells on pHH3-stained samples were annotated and counted in 10 HPFs as well.

2.6. Comparison of different methods

We compared the performance of pHH3 and Ki-67 immunohistochemistry visually and digitally in NST breast carcinomas with traditional mitotic count by HE-staining, assessing changes in histological grading, agreement, and correlation between methods. Mitotic indices on HE-stained slides and pHH3-stained slides were manually annotated and counted on 10 high power fields at 40x magnification, after being digitized. The visual assessment of Ki-67 positivity was based on the percentage of positively stained nuclei, digital assessment of

Ki-67 and pHH3 was performed using the Ki-67 quantifying tool of the CogM software, which automatically counted the total number of positive cells in each histological image.

The inter-rater reliability of different techniques for assessing mitotic figures was measured using intraclass correlation. The intraclass correlation coefficient (ICC) is a statistical measure to assess the reliability of a technique assessed by different observers and ranges from 0 to 1. It quantifies the degree of agreement among multiple raters and thereby indicates the extent to which the variability in data may be attributed to genuine differences between subjects.

Spearman correlation coefficient (SCC) measures the strength, direction, and the monotonic relationship between two variables. It captures the degree to which the variables move together in a consistent manner, either increasing or decreasing. In this study, the SCC was applied to evaluate the correlation between visual and digital assessments.

2.7. Cluster analysis

Cluster analysis was employed to establish new cut-off thresholds for pHH3, and Ki-67. The primary aim was to examine the interrelationship of visually and digitally assessed parameters, evaluate their similarities, and ascertain whether grouping them could yield distinct categories.

K-medoid cluster analysis is a widely used statistical method that aims to divide data points into k homogeneous groups or clusters. The method involves defining a set number of k clusters, each distinct from one another, but within themselves similar to each other. The value of k refers to the number of centroids, which represent the center of the clusters. The number of clusters was conducted multiple times, determining the optimal number k=3. The initial three cluster centers were randomly assigned to a scatter plot, every data point was allocated to the nearest centroid. The data was then averaged through repetitive calculations, leading to the identification of new centroids, and optimizing the clusters until the centroids have stabilized.

2.8. Statistical analysis

Categorical variables were described using absolute and relative frequency. Continuous variables were described as means and standard deviation or medians, and interquartile ranges (IQR) according to the distribution of the data. Normality of the distribution was assessed by investigating skewness and kurtosis as well as Q-Q graphs, box plots, and histograms. Correlations were calculated by Spearman´s rank order correlation since the data were not normally distributed. Furthermore, Fleiss Kappa was also performed to measure intra-rater agreement.

Agreement of continuous variables was assessed with intraclass correlation coefficient by means of a two-way mixed effects model.

A cluster analysis was conducted to identify groups of patients with as many similarities as possible. Previously, a standardization of variables was performed to get a mean of zero and standard deviation of one. The calculation of the distance between two real-valued vectors was performed using the Euclidean distance. PAM algorithm of K-Medoids clustering was the method used to group the patients based on their similarity. The Partitioning around Medoids (PAM) algorithm is a robust algorithm, which searches for K representative objects amongst a group of patients based on their similarity (k-medoids). In this method, each data point is assigned to the closest medoid, thereby creating clusters (objects within a cluster for which the average dissimilarity between it and all the other the members of the cluster are minimal). This algorithm was used since it is less sensitive to noise and outliers [89]. The clustering analysis was based on the following variants: histology classification (grade, mitosis score), immunohistochemistry (HE, Ki-67, pHH3) and visual vs. digital assessment. To estimate the optimal number of clusters, the average silhouette method was used (3 clusters were the number of clusters that maximized the average silhouette).

Statistical analysis was performed using IBM SPSS Statistics (SPSS Inc., Chicago, Illinois, USA) and R Software Version 4.0 (R Foundation for Statistical Computing, Vienna, Austria).

3. Results

3.1. Patient characteristics

The patient cohort included n=74 women, the median age was 58 (standard deviation: ± 14.9). Each case contributed 1-2 slides, resulting in a dataset of 84 slides stained with HE. This study included multiple slides of the same tumor for some cases, when available, to enhance the range of histologic features and tumor heterogeneity observed in real clinical cases. Additionally, these slides were included to ensure the dataset available for the deep learning algorithm was as extensive as possible.

Table 7: Patient characteristics: All cases with diagnosed NST breast cancer including gender, age histological grade, molecular subtype, and TNM (Tumor, Node, Metastases)-stage. Out of the total 74 cases, only 56 cases had available TNM stages. This limitation is attributed to the fact that not all treatment was carried out at the University of Tübingen.

Characteristics	**N=74**
Age (years)	
median	58
range	27-96
mean	61.2
SD	14.9
Gender	
women, n (%)	74 (100)
Histological grade (Elston and Ellis Classification)	
1, n (%)	10 (13,5)
2, n (%)	15 (20.3)
3, n (%)	49 (66.2)
Molecular subtype	
Luminal A, n (%)	13 (17.6)
Luminal B HER2+, n (%)	13 (17.6)
Luminal B HER2-, n (%)	16 (21.6)
HER2 positive, n (%)	18 (24.3)
Triple negative, n (%)	14 (18.9)
T stage, n (%)	N=56
pT0	15 (26,8)
pT1	21 (37,5)
pT2-4	19 (33,9)
N stage, n (%)	
Positive (pN1a-pN2c)	15 (26,8)
Negative (pN0)	41 (73,2)

3.2. Agreement between pathologists

The pathology report data from 84 HE slides were compared to the pathological re-evaluation conducted on the same slides. The comparison focused on the grade assessment according to the Elston and Ellis Grade, considering nuclear pleomorphism, tubular formation, and mitotic count.

There was an overall statistically significant interrater agreement ($p<0,001$) of 78,9% (95% CI 63,7% to 96,5%) between the two evaluations. Of the subgroups, mitotic score agreement was highest with 83,0% (95%CI 67,4 to 98,6), followed by tubular formation with 69,8% (95%CI 50,9% to 88,8%). Nuclear pleomorphism showed the poorest agreement with 62,9% (95% CI 42,0% to 83,7%). The Fleiss Kappa for grade I nuclear pleomorphism could not be calculated since there was only one case.

The correlation suggests that the assessments and evaluations of the histological features, such as nuclear pleomorphism, tubular formation, and mitotic count, were in strong alignment. Furthermore, with increasing histologic grade, the agreement between observers increased.

Table 8: Pathologists' interrater agreement (Fleiss Kappa)

	Fleiss Kappa	**95% CI**	**p-value**
Tubular formation	69,8%	50,9% – 88,8%	<0,001
score 1	48,8%	27,4% – 70,2%	
score 2	64,7%	43,4% – 86,1%	
score 3	71,4%	56,6% – 99,3%	
Nuclear pleomorphism	62,9%	42,0% – 83,7%	<0,001
score 1	-	-	
score 2	70,3%	40,5% – 83,3%	
score 3	92,3%	44,6% – 87,4%	
Mitotic index	83,0%	67,4% – 98,6%	<0,001
score 1	87,7%	60,0% – 100%	
score 2	89,7%	59,5% – 100%	
score 3	90,9%	67,3% – 100%	
Elson and Ellis grade	78,9%	63,7% – 96,5%	<0,001
grade I	77,8%	53,7% – 96,5%	
grade II	78,0%	49,6% – 92,3%	
grade III	95,4%	65,6% – 100%	

There was a very strong correlation (intraclass correlation coefficient (ICC) = 0.939; 95% CI 0.906 to 0.961; p<0.001) between two pathologists (pathology report vs. pathology re-evaluation) regarding the mitotic count according to the Elston and Ellis Classification.

3.3. HE mitotic count vs. Ki-67 proliferation index

The ICC for agreement between mitotic count on HE-stained slides and visual evaluation of Ki-67 IHC was 0.433 (95% CI: -0.167 to 0.712; Table 9), indicating moderate agreement. The interrater reliability was found to be lower when

comparing the results of mitotic count on HE-stained slides obtained via visual assessment (VA) with those of Ki-67 IHC obtained via digital assessment (DA) (ICC= 0.279; 95% CI: -0.156 to 0.561; Table 9).

Table 9: ICC of VA of mitotic count (HE) vs. VA of Ki-67 and VA of mitotic count (HE) vs. DA of Ki-67

	ICC (95% CI)	p-value
VA of mitotic count (HE) vs. VA of Ki-i67	0.433 (-0.167 to 0.712)	<0.001
VA of mitotic count (HE) vs. DA of Ki-67	0.279 (-0.156 to 0.561)	<0.002

A moderate correlation was identified between the visual evaluation of mitotic count on HE-stained slides and the visual assessment of Ki-67 staining, when applying the Spearman correlation. The strength of the correlation was quantified by a SCC of 0.525, indicating a moderate correlation between the two techniques (Table 10). Additionally, a moderate correlation was observed between the visual assessment of mitotic count on HE-stained slides and the digital assessment of Ki-67-stained slides, with a SCC of 0.434 (Table 10).

Table 10: Spearman correlations between mitotic count on HE stains vs. VA of Ki-67 and mitotic count on HE stains vs. DA of Ki-67 within histological grades

	Spearman correlation
Mitotic count (HE) vs. Ki-67 visual assessment	0.525
Mitotic count (HE) vs. Ki-67 digital assessment	0.434

Upon analyzing the correlation within individual histological grades (as shown in Table 11), a strong positive correlation was found between the visual and digital comparison of Ki-67 for histological grades II and III, with SCC of 0.806 and 0.728, respectively. A moderate correlation, represented by a SCC of 0.563, was

observed for grade I. By comparing the VA and DA of Ki-67 within each histological grade, it allows for the evaluation of the agreement between these two methods at the grade level.

Table 11: Spearman correlations between VA and DA of Ki-67 within histological grades

Histological grade	Spearman correlation
grade I	0.563
grade II	0.806
grade III	0.728

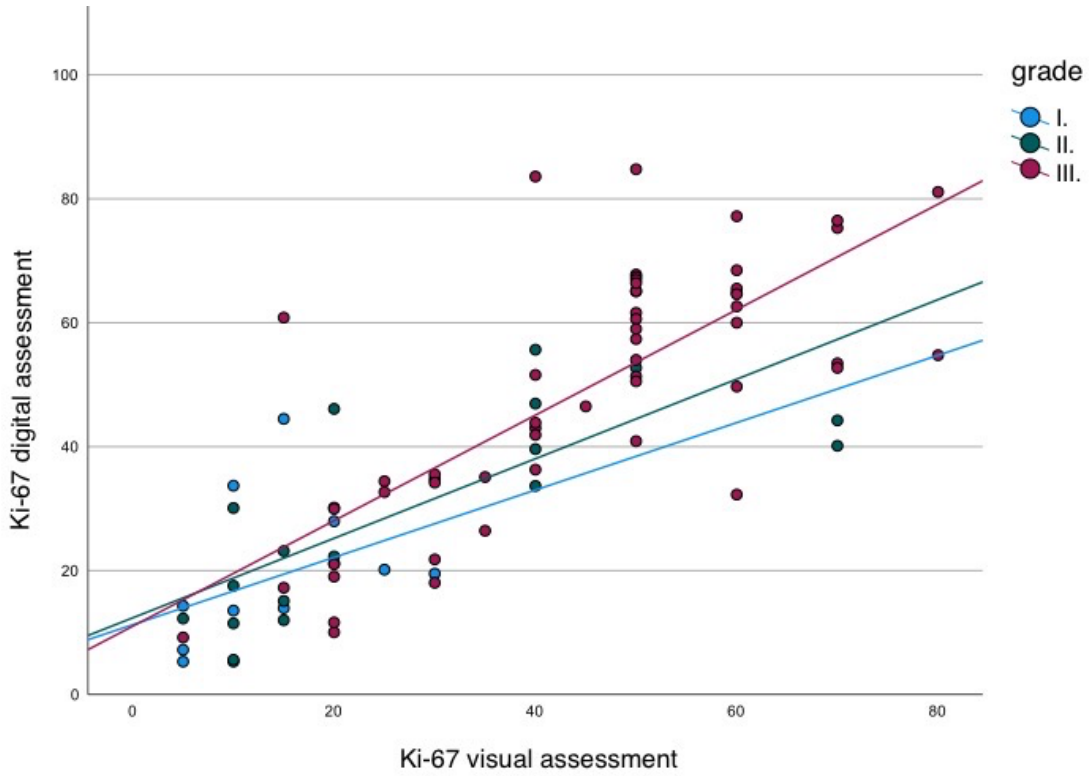


Figure 7: Comparison between visual and digital assessment of Ki-67, according to the original histological evaluation (grade I, II, III). Scatter plots with regression lines. VA (x-axis) is plotted against DA (y-axis).

Figure 7 depicts the correlation between VA and DA of Ki-67. A positive and linear relationship between the VA and DA of Ki-67 is observed, with tightly clustered

data points around the regression lines, suggesting a strong relationship between the two methods.

3.4. HE mitotic count vs. pHH3 mitotic indices

Table 12 presents the number of slides with mitosis scores from 1 to 3, assessed by different methods, including HE visual, pHH3 visual, and pHH3 digital, categorized by their respective histological grade.

According to VA of HE-staining, 18 out of 84 slides (21.4%) were categorized as having a mitosis score of 3. The pHH3 marker revealed a greater number of mitotic cells, thus in the visual analysis this number increased to 72 out of 84 slides (85.7%), and further to 100% when DA was used. However, this increase diminished the discriminatory properties, as all cases were then categorized as having a mitosis score of 3.

Figure 8 depicts the distribution of the number of mitoses based on mitosis scores 1-3, as determined by various methods (HE visual, pHH3 visual, and pHH3 digital).

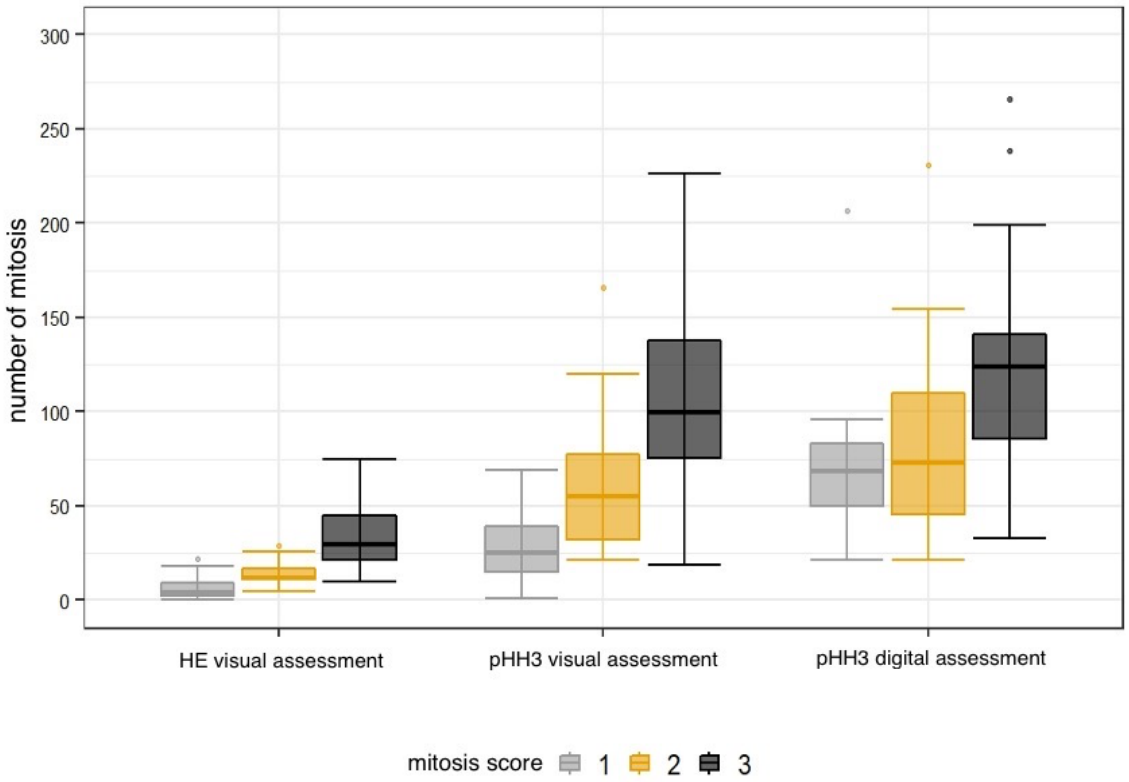


Figure 8: Number of mitoses (y-axis) plotted against different methods of analysis (x-axis: Visual assessment of HE-stained slides; visual and digital assessment of pHH3). Further subdivision of each analysis method into mitosis score 1 (gray), 2 (yellow) and 3 (black).

Table 12: Stratification of mitosis scores as assessed by different methods (HE visual, pHH3 visual, and pHH3 digital).

Grade	VA of HE-stained slides (n, %)	VA of pHH3-stained slides (n, %)	DA of pHH3-stained slides (n, %)
Score 1	29 (34.5%)	4 (4.8%)	0
Score 2	37 (44.0%)	8 (9.5%)	0
Score 3	18 (21.4%)	72 (85.7%)	84 (100%)
Total	84	84	84

The mitosis count is scored by different assessment methods (see 3 columns) and the number of slides of each scoring category as well as the percentages are

depicted (total n=84 slides). The number of slides according to mitosis score was assessed according to mitotic count/area (0.159mm^2; Table 1).

3.4.1. Agreement between methods

The study compared the initial assessment of mitotic count on HE-stained slides with the evaluation of mitotic indices on pHH3-stained slides, using both visual and digital assessment techniques. The obtained intraclass correlation coefficients were 0.315 (with a 95% confidence interval ranging from -0.156 to 0.599; $p < 0.001$, as shown in Table 13) for comparing mitotic count on HE-stained slides with visual assessment of pHH3, and 0.134 (with a 95% confidence interval ranging from -0.131 to 0.378; $p = 0.039$, as shown in Table 13) for comparing mitotic count on HE-stained slides with digital assessment of pHH3. These results suggest that the assessment of the two different stains may not be interchangeable techniques. A high degree of reliability was found between VA and DA of pHH3. The average ICC measure was 0,849 (with a 95% CI from 0.365 to 0.941; p<0.001, Table 13).

Table 13: ICC of mitotic count on HE-stained slides and VA of pHH3-stained slides, mitotic count on HE-stained slides and DA of pHH3-stained slides and between VA and DA of pHH3-stained slides

	ICC (95% CI)	p-value
VA of mitotic count (HE) vs. VA of pHH3	0.315 (-0.156 to 0.599)	<0.001
VA of mitotic count (HE) vs. DA of pHH3	0.134 (-0.131 to 0.378)	0.039
VA of pHH3 vs. DA of pHH3	0.849 (0.365 to 0.941)	<0.001

A moderate correlation was found between mitotic count on HE-stained slides (assessed visually) and the number of pHH3 positive cells (assessed visually) indicated by the SCC of 0.643, see Table 14.
However, a weak correlation was observed when comparing the mitotic count on HE-stained slides (assessed visually) with the number of pHH3 positive cells (assessed digitally), as indicated by the SCC of 0.384.

Table 14: SCC of mitotic count on HE slides and pHH3 positive cells by VA and DA

	Spearman correlation
Mitotic count (HE) vs. pHH3 VA	0.643
Mitotic count (HE) vs. pHH3 DA	0.384

Upon analyzing the Spearman correlation within individual histological grades, a very strong positive correlation between visual and digital comparison of pHH3 across all three histological grades was observed (Table 15), as indicated by the SCC of 0.955 for grade I, 0.763 for grade II, and 0.889 for grade III.

Table 15: Spearman correlations between VA and DA of pHH3 within histological grades

Histological grade	**Spearman correlation**
grade I	0.955
grade II	0.763
grade III	0.889

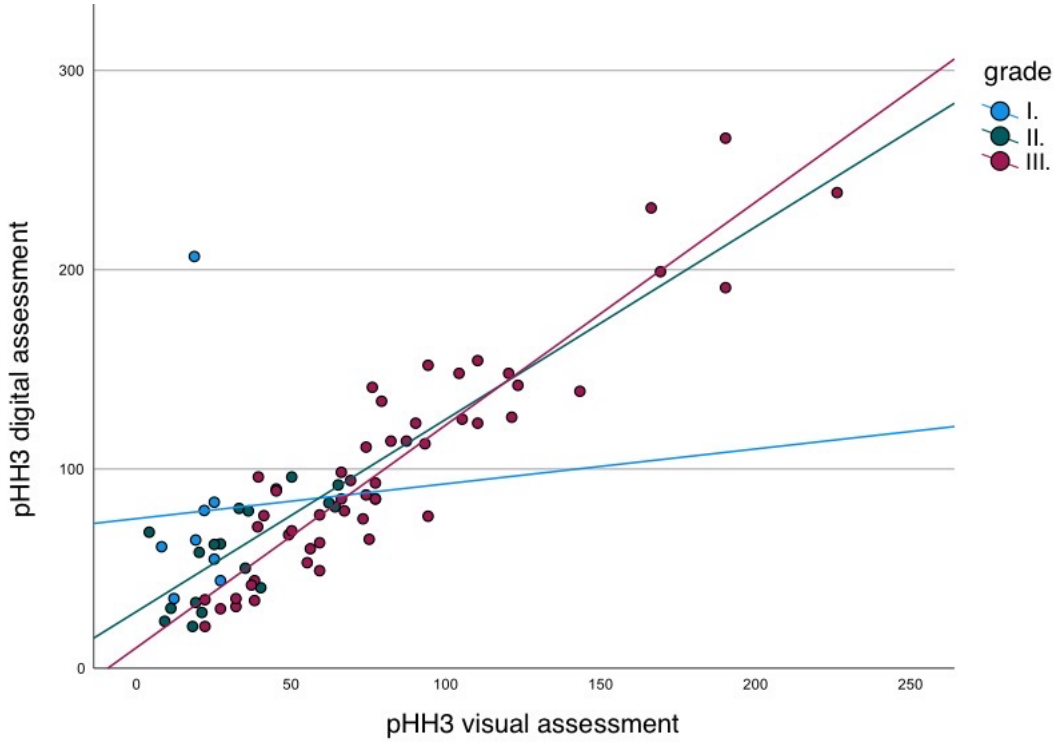


Figure 9: Comparison between visual and digital assessment of pHH3 according to the original histological evaluation (grade I, II, III). Scatter plots with regression lines. VA (x-axis) is plotted against DA (y-axis).

Figure 9 illustrates the positive and linear relationship between the VA and DA of pHH3, where the x-axis and y-axis represent the VA and DA, respectively. The dots in the plot represent the individual data points, with different colors representing the histological grades I-III.

In particular, the purple data points representing grade III are tightly clustered around the purple regression line, indicating a strong and consistent relationship between VA and Da within this grade.

3.5. PHH3 mitotic indices vs. Ki-67 proliferation index

The ICC for agreement between visual evaluation of positive cells on Ki-67-stained slides and visual evaluation of positive cells on pHH3 stain was 0.517 (95% CI: 0.172 to 0.709; <0.001), indicating moderate agreement.

Table 16: ICC of VA of pHH3 vs. VA of Ki-67 and of DA of pHH3 vs. DA of Ki-67

	ICC (95% CI)	p-value
VA of pHH3 vs. VA of Ki-67	0.517 (0.172- 0.709)	<0.001
DA of pHH3 vs. DA of Ki-67	0.324 (-0.131 - 0.598)	<0.002

The ICC for agreement between the digital evaluation of positive cells on Ki-67-stained slides and digital evaluation of positive cells on pHH3 stain was 0.324 (95% CI: -0.131 to 0.598; p<0.002), indicating moderate agreement.
A moderate correlation was observed between visual and digital assessment of Ki-67 stain (SCC = 0.668), as determined by the Spearman correlation. A slightly lower SCC of 0.478 was found, when comparing positive cells on pHH3 and Ki-67-stained slides evaluated by DA.

Table 17: Spearman correlation between VA and DA of pHH3 and Ki-67 IHC

	Spearman correlation
pHH3 visual vs. Ki-67 VA	0.668
pHH3 digital vs. Ki-67 DA	0.478

3.6 Identification of new thresholds based on quantitative assessment

The objective of utilizing cluster analysis was to explore the relationship between different parameters, assess their similarities, and determine if grouping them together would result in distinct categories. This analysis seeks to identify patterns among the parameters, enabling the identification of clusters that share common characteristics or properties. Utilizing cluster analysis provides insights

into the underlying structure and organization of the data, which can aid in classifying different categories based on the similarities observed among the parameters.

3.5.1. Cluster analysis 1

To identify new cut-off thresholds based on quantitative assessment of pHH3 and Ki-67, cluster analysis was performed, with the optimal number of clusters determined to be three. The results are depicted in Figure 10, which illustrates a scatter plot of visual quantification of pHH3 against visual quantification of Ki-67, with the plot divided into three clusters representing potential cut-off values. Each point in the scatter plot represents the number of mitoses for a single tissue sample. The data points are colored and grouped into three distinct clusters (1-3), based on the results of the k-medoid cluster analysis. These clusters represent possible cut-off values for the number of mitoses. The scatter plot allows for visualization of the relationship between pHH3 and Ki-67 staining and their association with the number of mitoses in each tissue sample.

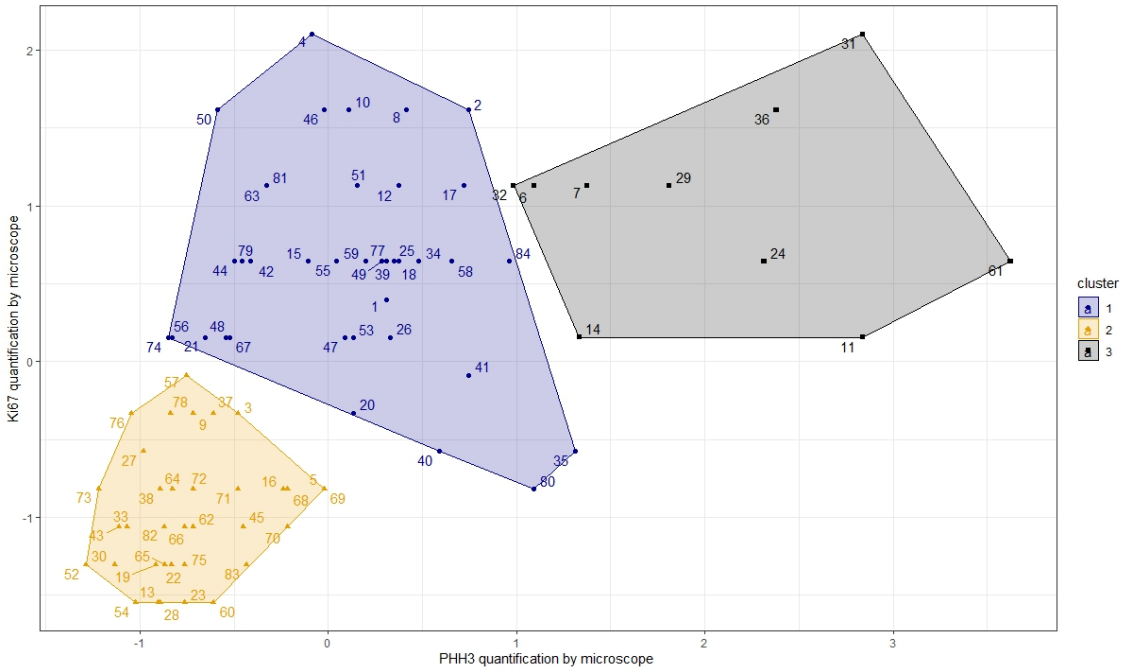


Figure 10: Distribution of number of mitoses according to visual quantification of pHH3 (x-axis) and visual quantification of Ki-67 (y-axis). Clusters 1-3 are color-coded

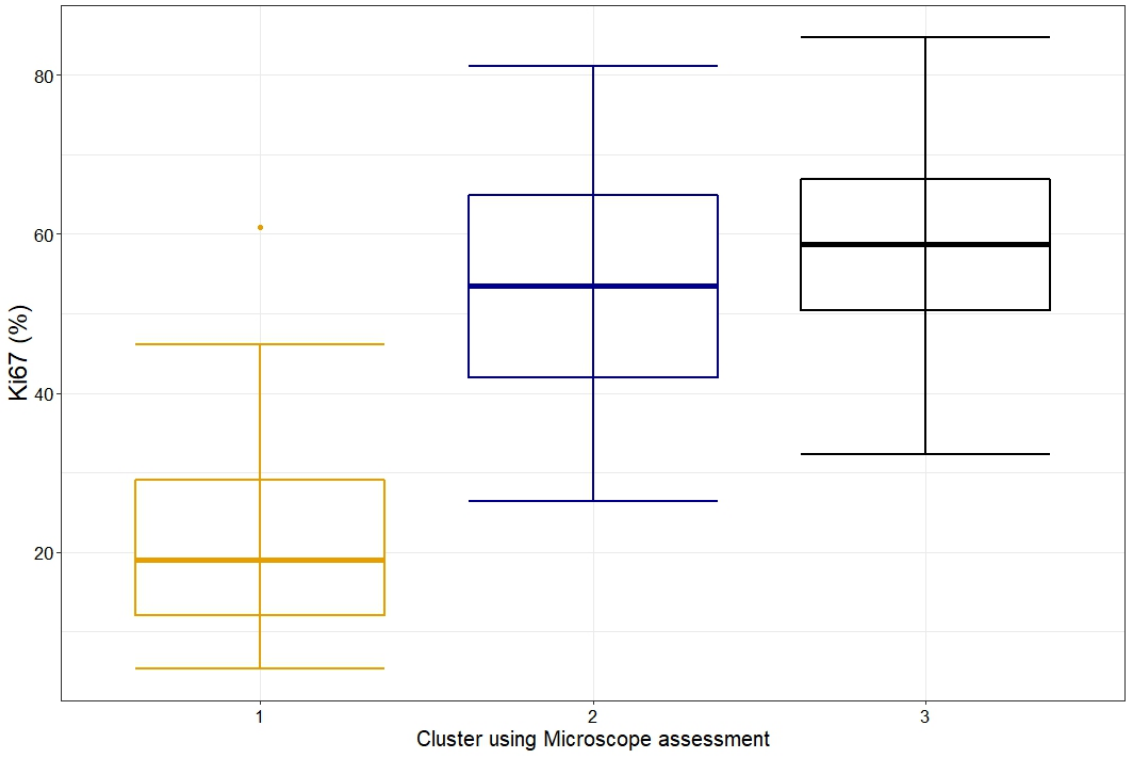


Figure 11: Visualization of percentage of Ki-67 according to clustering 1

In Figure 11, boxplots were constructed to represent the distribution of Ki-67 according to the clusters obtained from cluster analysis 1. The overlap in interquartile ranges between group 2 and 3 suggests that the values within this range could be classified into both groups, leading to reduced specificity. In contrast, Figure 12 presents a boxplot illustrating the classification of mitotic indices on pHH3 IHC, according to the clusters derived from cluster analysis 1. The interquartile ranges of the clusters do not overlap, indicating that the analysis of cluster 1 is more suitable for establishing cut-off values for pHH3.

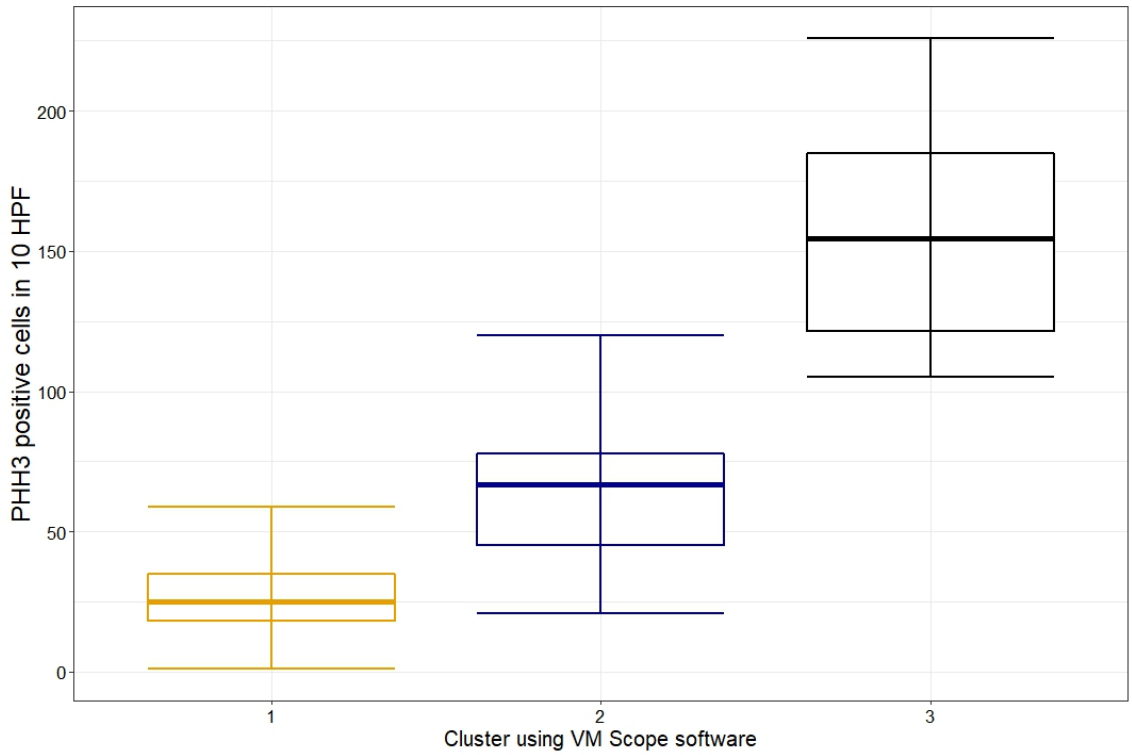


Figure 12*:* Visualization of pHH3 positive cells according to clustering 1

3.5.2. Cluster analysis 2

Figure 13 displays a scatter plot representing the relationship between visual quantification of pHH3 (x-axis) and digital quantification of Ki-67 (y-axis), with clusters identified through a similar procedure as described in cluster analysis 1. The scatter plot is divided into three clusters, with each cluster potentially representing a different cut-off value for pHH3 and Ki-67.

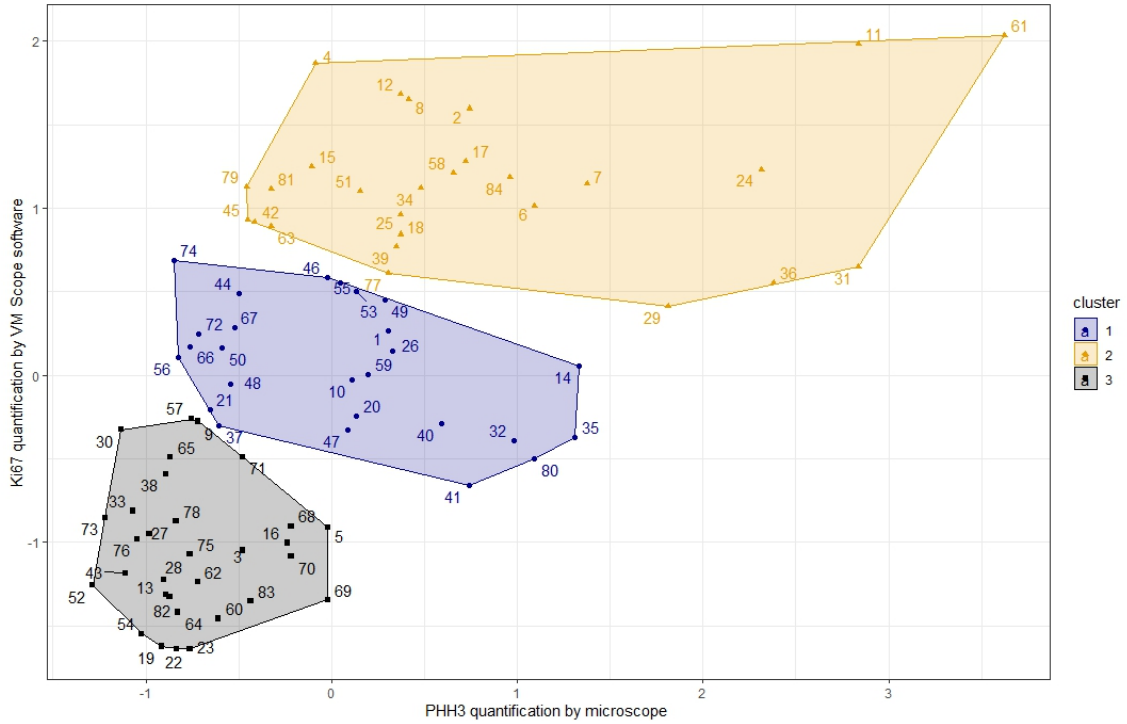


Figure 13: Distribution of number of mitoses according to clusters 1-3, according to visual quantification of pHH3 (x-axis) and digital quantification of Ki-67 (y-axis).

In Figure 14, the distribution of Ki-67 according to cluster 2 is displayed using boxplots. The interquartile ranges do not overlap, indicating a higher specificity for creating cut-off values. In contrast, Figure 15 shows the distribution of pHH3 according to cluster analysis 2, which displays an overlap of interquartile ranges in groups 2 and 3. This overlap signifies a higher probability for categorizing values into both groups, indicating a lower specificity for creating cut-off values for pHH3.

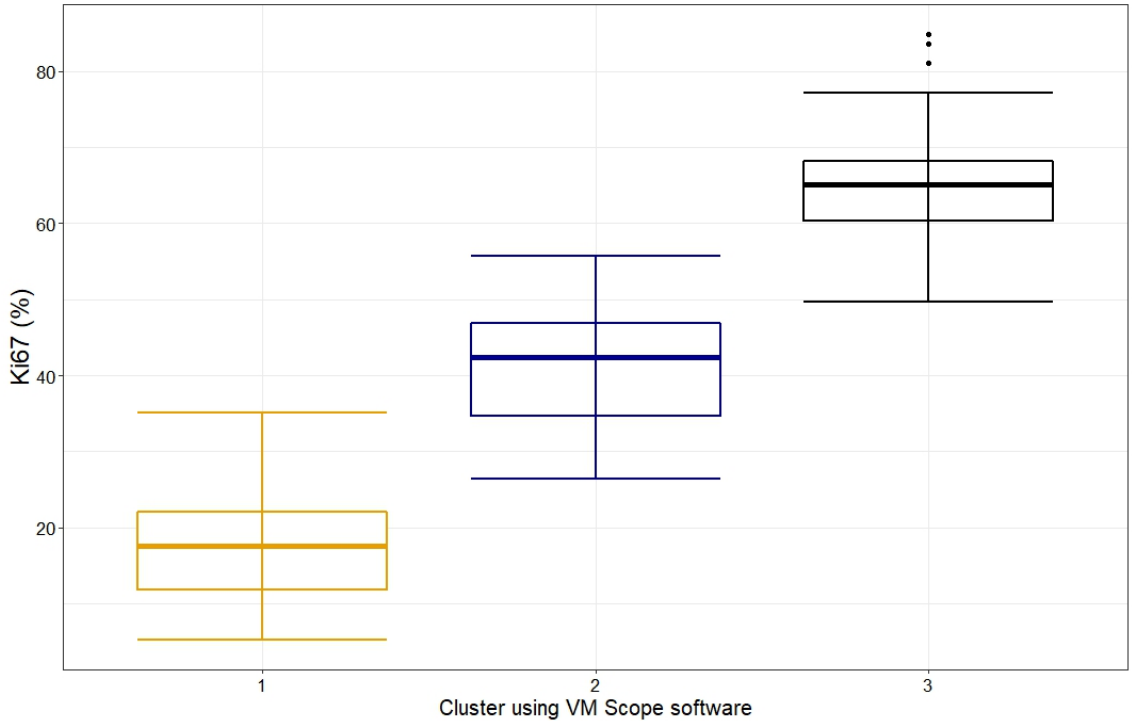


Figure 14: Visualization of percentage of Ki-67 according to clustering 2

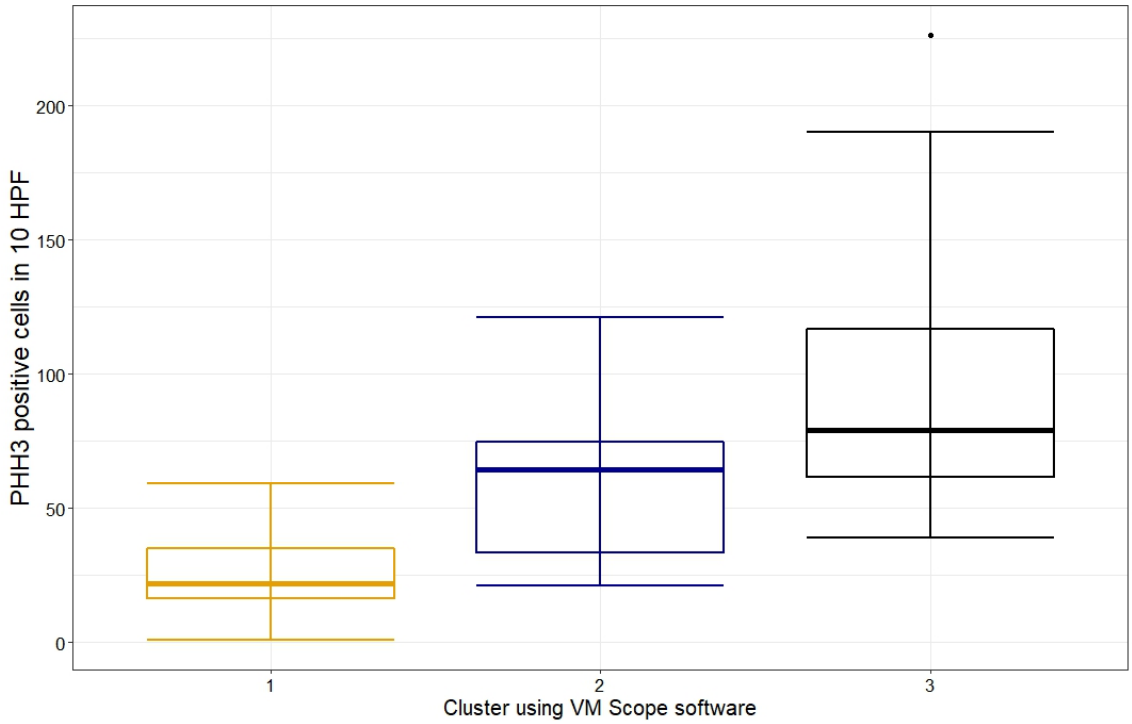


Figure 15: Visualization of pHH3 positive cells according to clustering 2

Table 18: Min-max values of possible groups stratified by clustering 2

Grade	1	2	3
pHH3 Visual assessment	1 - 59	21 - 121	39 – 226
Ki-67 (%) Visual assessment	5 - 35	15 – 70	15 – 80
Ki-67 (%) Digital assessment	5.3 - 35	26.4 – 55.7	49.7 – 84.8

New possible subgroups could be identified through the application of cluster analysis, based on the quantitative assessment of pHH3 and Ki-67. The number of positive cells on pHH3 IHC and percentages of proliferating cells on Ki-67 IHC were assigned to three possible groups, grade I to III, according to the second clustering.

3.6. Summary of the results

The study examined the impact on histological grading, agreement, and correlation between methods. We used the ICC to assess histological slides with different methods, the SCC to determine the relationship between techniques and finally, through cluster analysis we propose potential cut-off values for the biomarkers pHH3 and Ki-67.

The VA of pHH3 showed higher sensitivity in detecting mitotic figures compared to mitotic count in HE-stained slides, suggesting that these approaches are not equivalent. The agreement between mitotic count and the VA of both immunohistochemical methods, as well as between the two immunohistochemical methods pHH3 and Ki-67 was moderate (ICC= 0.315 to 0.517; SCC=0.525 to 0.668).

However, considering pHH3 immunolabeling, VA allows for better identification of the area with the highest mitotic activity due to the distinct contrast between stained elements.

The agreement between mitotic count and DA was comparatively lower (ICC=0.134 to 0.279; SCC=0.384 to 0.434).

The strong correlation observed between the visual and digital approaches indicates the potential of the digital method as a viable alternative to conventional counting methods. Despite the superior performance of digital pathology in quantifying Ki-67 IHC staining, our study demonstrated that it was not feasible for accurate quantification of pHH3 due to the software's limited ability to distinguish between dark artifacts and positive cells. Moreover, to identify new cut-off thresholds based on quantitative assessment of pHH3 and Ki-67, cluster analysis was performed, with the optimal number of clusters determined to be three. Accordingly, the study proposes the use of 3 subgroups for pHH3 with minimum to maximum values (such as group 1: 1 – 59; group 2: 21 –121; group 3: 39-226) and 3 subgroups for Ki-67 (such as group 1: 5,3 - 35 %; group 2: 26,4 –55,7 %; group 3: 49,7 - 84,8 %).

4. Discussion

In this study, the main objective was to evaluate the accuracy and interobserver variability of the Elston and Ellis Classification in NST breast carcinoma and to compare the conventional HE-staining with Ki-67 and pHH3 as quantifiable immunohistochemical markers. Another goal was to make the histological grading process more objective by utilizing digital pathology for quantitative measurements.

The results of our study indicate that the pHH3 marker offers a higher degree of sensitivity and accuracy in determining mitotic count than the conventional HE-staining. Our findings highlight the potential utility of incorporating pHH3 as a quantifiable marker in the process of histological grading for more objective differentiation, particularly in challenging cases between grades II and III. However, the establishment of pHH3-specific cut-off values would be required.

Despite the superior performance of digital pathology in quantifying Ki-67 immunohistochemical staining, our study demonstrated that it was not feasible for accurate quantification of pHH3 due to the software's limited ability to distinguish between dark artifacts and positive cells.

In the following chapters these results will be discussed further.

4.1. Accuracy and interobserver variability of the Elston and Ellis Score

A large-scale study revealed that the mitotic count was the only factor that had a significant impact on prognosis, when individual elements of the Elston and Ellis Classification were examined[90]. Moreover, numerous studies have demonstrated its independent prognostic value[91,92].

The current method for determining the mitotically active cells is identifying them based on their morphology on HE-stained slides. However, the identification of mitotic figures in practical application may be challenging, as it has been shown to be highly subjective[93,94] and prone to considerable inter- and intra-observer variability. This variability may arise due to variations in the interpretation of

morphological features, differences in identifying mitotically active areas on HE-stained slides, as well as variability in slide fixation and thickness. The presence of cells resembling mitotic figures such as apoptotic, karyorrhectic or hyperchromatic cells may further diminish the reproducibility[65].

Additionally, cells in metaphase and anaphase are typically easier to identify, while recognizing cells in prophase or telophase can be difficult or even impossible. Moreover, crushed artifacts, which often obscure morphological details, may further hinder the accurate identification of mitotic figures. These factors may weaken the reliability of histologic grade, particularly when mitotic count is a crucial factor in determining the overall grade, especially when the score falls within the intermediate range based on tubular formation and nuclear pleomorphism.

In this study, two pathologists evaluated a total of 84 HE-stained slides. The interobserver agreement was 83.0% for mitotic count, 69.8% for tubular formation, and the lowest agreement of 62.9% for nuclear pleomorphism. The overall agreement between the two observers was 78.9%. These findings align with a previous study conducted by Elmore et al., which reported an interobserver agreement of approximately 75% among three consensus panel members for breast biopsy diagnoses[95]. Many other studies have been published regarding the variability in pathologists´ breast cancer grading, with a wide range of agreement[94,96-99]. While some studies have reported lower agreement for nuclear pleomorphism[94,99], as observed in our study, others have found greater variability in mitotic count[98,100]. Our study also revealed that as the histologic grade increased, the agreement between observers also increased.

In contrast, a study by Bueno-de-Mesquita et al. found that inter-observer variation was higher in grade II tumors than in grade I and III tumors[101]. Interestingly, there was a very strong correlation (ICC= 0,928) between the two pathologists regarding the Elston and Ellis Classification, despite the commonly reported subjectivity and moderate interobserver reproducibility associated with mitotic count assessment[17,66]. It is noteworthy, that this strong correlation could be due to both pathologists working at the same institute and the relatively low sample size of the study.

4.2. Comparison of HE mitotic count and Ki-67 proliferation index

Ki-67 is an indicator of proliferation and a clinically validated predictor of prognosis for early breast cancer[57,59] and a prognostic factor for several tumors including cutaneous melanomas[102] and meningiomas[103].

Despite the widespread use of Ki-67 IHC for estimation of cell proliferation, its utility in both breast cancer research and clinical decision-making has come into question due to the considerable variability in Ki-67 results. Determining the ideal cut-off value for Ki-67 in clinical practice has proven difficult, further complicating its utility. The published cut-off for Ki-67 in different studies vary between 5% and 34%[104] and the International Ki-67 in Breast Cancer Working Group (IKWG) was unable to come to a consensus regarding the ideal cut-offs that might be used in clinical practice[31]. In a 2009 study by Cheang et al., which was the first study to attempt to distinguish the luminal B subtype from the luminal A subtype, a cut-off of 14% was proposed[58]. In 2011 this threshold was suggested by the St. Gallen Consensus Meeting[69] and was upgraded to 20% two years later[95].

In a meta-analysis conducted by Azambuja et al., a wide range of Ki-67 cut-off values were found across different studies[57], which included a 10% threshold, mean or median values. Therefore, these differences might contribute to the difficulty in setting a standard cut-off value in daily practice. The differences in cut-off values may also be influenced by the clinical objective, as suggested by some authors[57,105]. For instance, a 10% cut-off may be useful in cases with slow proliferating tumors to avoid overtreatment, while a 25% cut-off may be appropriate for identifying patients who could benefit from chemotherapy protocols, as demonstrated in a study by Spyratos et al.[57,105].

Denkert et al. proposed that Ki-67 should be regarded as a continuous variable as it reflects the percentage of proliferating cells in the tumor, which can range from 0% to 100%. They suggested that setting a standard cut-off value might not be possible, and that Ki-67 levels should be regarded as dependent on clinical endpoints, treatment, and molecular subtypes[106]. However, a continuous marker may offer more flexibility for clinical decision-making and the ability to adjust to

different therapeutic strategies. IKWG states that Ki-67 levels less than 5% and 30% or more should be considered clinically actionable[29].

We chose to investigate the relationship between Ki-67 and mitotic count since previous research has shown a strong association between Ki-67 and tumor grade[107,108], which is an important prognostic factor in breast cancer[5]. Additionally, mitotic count is an important component of histologic grade, therefore, evaluating the correlation between these two measures was of interest in this study.

In the present study, the agreement between mitotic count on HE-stained slides and visual evaluation of Ki-67 IHC was determined by intraclass correlation coefficient analysis. The agreement between mitotic count on HE-stained slides and visual evaluation of Ki-67 IHC was moderate agreement (0.433; 95% CI: -0.167 to 0.712). Furthermore, the ICC agreement between mitotic count on HE-stained slides and digital evaluation Ki-67 IHC was also assessed, revealing slightly lower agreement (ICC= 0.279; 95% CI: -0.156 to 0.561). Given that Ki-67 can be detected in all phases of the cell cycle except for the resting phase, while mitotic figures in HE-staining are only visible in the mitotic phase, it is not unexpected that there is only moderate correlation between these two measures. Therefore, they may provide complementary information, but cannot be used interchangeably.

Furthermore, when comparing the mitotic count with visual and digital assessment of Ki-67 IHC, the SCC indicated a moderate correlation (mitotic count (HE) vs. Ki-67 VA = 0.525; mitotic count (HE) vs. Ki-67 DA = 0.434). The moderate correlation could be attributed to the inherent differences in the measurement approaches of the two methods to assess cell growth and division. While mitotic count reflects the number of mitotic figures identified during the mitotic phase of the cell cycle, Ki-67 is a marker of proliferation, present during all active phases of the cell cycle except the resting phase. The potential influence of molecular subtypes on the correlation between mitotic count and Ki-67 IHC was not considered in the present study. It is possible that differences in the molecular subtypes of breast cancer may have an impact on the observed correlation between these two methods of measurement[55,109].

In conclusion, the moderate correlation observed between the mitotic count and Ki-67 IHC in this study may be attributed to multiple factors, such as the differences in methodology of the two techniques, as well as potential variations in the molecular subtypes of the tumors.

4.2.1. Comparison of visual and digital assessment of Ki-67 scoring

In this study, two methods were employed for the evaluation of Ki-67 in breast cancer: visual assessment and automated digital analysis.

Digital pathology is a promising tool in reducing interobserver variability, increasing accuracy and reproducibility in diagnostic and research settings[110]. However, until recently, the image quality and analysis software did not allow for automated assessment of histological slides. According to Robertson et al. the ultimate goal of digital pathology is to improve pathologists' workflow, provide more accuracy, reduce human error, and increase reproducibility[33].

Although many image analysis systems exist, most provide the opportunity to adjust parameters, which may lead to inconsistent results. The non-adjustable software of Cognition Master Professional Suite (CogM), a collection of image analysis software tools, offers an advantage for routine application, as no opportunity exists for modification. CogM enables the presentation, evaluation, and analysis of digital histological slides, making it a valuable tool for digital pathology. The Ki-67 Quantifier module used for this study has been previously validated in a neoadjuvant breast cancer clinical trial as a computer-based approach for Ki-67 scoring[34]. A recent study conducted by Alataki et. al., demonstrated that CogM image analyzes allowed for standardized automated Ki-67 scoring that accurately replicated previously clinically validated and calibrated manual scores[111].

Our analysis showed good agreement between the conventional (visual) and digital methods of Ki-67 expressions (SCC: grade I = 0.563; grade II = 0.806; grade III = 0.728). This is in line with previous publications that also obtained high levels of correlation[112,113]. Given the high correlation observed between the visual and digital methods of Ki-67 expression, the digital method shows promise as a

potential alternative to traditional counting methods, as digital pathology continues to advance.

However, the IKWG currently advises the manual counting of Ki-67 through light microscopy or from a digital image, while automated scoring is still being researched[29].

In their discussions, the IKWG also addressed different approaches to Ki-67 assessment, namely the hot spot score and the global score. The hot spot score involves selecting the field with the highest observed Ki-67 index and scoring up to 500 cells. To obtain the global score, the IKWG developed an online scoring application that evaluates 100 cells in four distinct areas of the tumor section (negligible, low, medium, or high), aiming to capture tumor heterogeneity. In studies conducted by the IKWG, global scores obtained across the entire tumor section demonstrated higher reproducibility compared to hot spot methods. However, it is important to note that these differences did not reach statistical significance, indicating the need for further advancements in this area[29].

4.2.2. The advantages of using pHH3 as a mitosis specific marker

The pHH3 marker is a specific marker for mitoses that offers a potentially more objective method for identifying mitotic figures compared to the traditional mitotic count. This approach has several advantages, including reducing the uncertainties and misdiagnoses due to artifacts associated with traditional methods and enabling the detection of easily missed mitoses.

Studies have shown that pHH3 is a practical, robust and sensitive indicator of mitotic figures in breast cancer, enabling faster and easier detection of mitoses and the identification of hot-spots[65]. Moreover, pHH3 expression has been demonstrated to be a stronger prognostic marker than axillary lymph node status, tumor size, nuclear grade, and histological grade[23].

Comparable to mitosis scoring on HE samples, mitotic indices on pHH3-stained slides involve counting the total number of tumor cells in 10 consecutive high-power fields (HPFs) where the immunolabeling was most prevalent. The use of pHH3 in identifying additional mitoses may have practical application. Recent studies have shown pHH3 IHC to be highly specific for phosphorylated histone

H3, which is always heavily phosphorylated during metaphase. Histone H3 is a nuclear core histone protein that is part of chromatin, and its phosphorylation at serine-10 and serine-28 likely plays an important role in chromosome condensation and cell-cycle progression during mitosis[114]. Especially metaphase chromosomes are heavily phosphorylated, while interphase cells do not stain or show very low intensity[114,115]. Hence, pHH3 counts should theoretically correlate with mitotic counts and several reports have been published on the positive correlation between mitotic count and pHH3 counts [22,24,116,117]. Research studies have indicated that pHH3 is a useful marker in several types of tumors. PHH3 IHC has been recommended as an adjunct to HE-staining, as it has been shown to improve inter-rater reliability in identifying mitotic “hot spots” in thin melanoma, thereby helping diagnosis[118]. Similarly, in well-differentiated neuroendocrine tumors of the pancreas, the use of pHH3 IHC has been shown to improve interobserver agreement in mitotic count and grade assessment compared to HE[64]. In a study by Skaland et al. pHH3 expression served as a strong prognostic marker for lymph node-negative breast cancer patients under the age of 55 years who underwent systemic adjuvant chemotherapy[22] and found that a threshold of ≥13 positive cells for pHH3 could effectively distinguish between low-proliferative and high-proliferative tumors, with the cut-off value of 13 providing the strongest prognostic threshold for 20 year recurrence-free survival[22]. In a study conducted by Van Steenhoven et al., the threshold of ≥13 positive cells for pHH3 was utilized and the findings demonstrated that pHH3 exhibits independent prognostic value, thus potentially improving the histological grading process by enabling more accurate identification of mitotic figures[119].

4.3. PHH3 as a quantifiable marker

The use of pHH3 staining for visual and digital analysis has shown a higher sensitivity in identifying mitotic figures compared to traditional HE-staining. The re-evaluation of mitosis score using pHH3 staining resulted in a significant upgrade of mitotic score and grade, with a higher number of mitotic figures being identified. According to visual assessment of HE-staining, 18 out of 84 slides (21.4%) were categorized as having a mitosis score of 3. However, using the

pHH3 marker in visual analysis increased this number to 72 out of 84 slides (85.7%), and further to 100%, when digital assessment was used.

The higher sensitivity of pHH3 staining was expected and is likely due to the fact that pHH3 staining allows for the identification of prophase figures that are impossible to count with regular HE staining, but can be easily identified in pHH3-stained slides[116]. Furthermore, pHH3 allows rapid detection of the mitotically most active area because of the sharp contrast between stained elements compared to non-stained elements. The increased sensitivity and accuracy of pHH3 staining have been demonstrated in several other studies concerning breast cancer[22,116,120]. A study by Veta et al. suggested that automated image analysis of pHH3 immunostains could be used to guide pathologists to the region of the tumor with the highest proliferation rate and help in improving objectivity[53].

The evaluation of pHH3-stained slides using automated digital analysis led to a mitosis score of 3 in all cases (100%), indicating a lack of discriminatory power when categorizing the number of positive cells into different grades according to the original HE-staining assessment. The recognition tool of the software used for cell segmentation may have been disturbed, particularly by darker artifacts. Segmentation of nuclei is a challenging task due to the complex shapes and overlapping of nuclei, imperfect slide preparation or staining, contact regions, and artifacts[8,121].

In their study “Grading of invasive breast carcinoma: the way forward”, Van Dooijeweert et.al. quoted an important remark by Bloom and Richardson form 1957 that the three histological groups are not different entities, but statistically allocated cut-off values[2,122]. As the current cut-off values for mitotic count are based on studies conducted using HE-staining, they may not be directly applicable to pHH3-staining. Consequently, it is necessary to establish specific cut-off values for pHH3-staining in the future.

The comparison of mitotic count based on HE-staining with the evaluation of pHH3 positive cells showed a stronger correlation when pHH3 was assessed visually, with a correlation coefficient of 0.643. However, digital assessment of pHH3 was not always reliable due to difficulties in distinguishing between artifacts, darker staining, and actual positive cells. As a result, interrater

agreement between mitotic count on HE and pHH3 positive cells with digital assessment was only moderate, with a correlation coefficient of 0.384. The observed increase in mitotic counts based on pHH3 staining in comparison to HE-staining could be attributed to the higher sensitivity and specificity of pHH3 in staining mitotic figures and highlighting nuclear details. Additionally, the positive cells on pHH3-stained slides are easier to distinguish than those on HE-stained slides, leading to a higher count. This finding is consistent with a study by Cui et al. which reported a similar increase in mitotic count based on pHH3 staining[66].

A strong correlation was observed between the visual and digital assessment of pHH3-stained slides, as indicated by a strong SCC and tightly clustered scatter plots around the regression lines in Figure 9, especially for grade III. This suggests a strong and linear relationship between the evaluation methods. These findings are consistent with a previous study by Bossard et al., which also reported a good correlation between visual and computer-assisted counts of pHH3 positive mitotic figures in breast adenocarcinomas[116].

Using a quantifiable biomarker such as pHH3 with acomputer assisted method might offer the opportunity to better reproducibility, and accuracy[24,123]. Although a high degree of reliability (ICC=0,849) was found between visual and digital assessment of pHH3, the digital method was not feasible due to technical limitations. Specifically, the software used for digital assessment (CogM) could not always distinguish between artifacts, darker staining, and actual positive cells, which led to inaccurate results.

4.4. Comparison of pHH3 mitotic indices and Ki-67 proliferation index

Ki-67 and pHH3 are both quantifiable IHC markers, which allow fast recognition and easy distinction between positive and negative cells. Both markers may be employed to detect and quantify cells in the process of mitosis. Subsequently, a comparison was conducted to evaluate positive indices of pHH3 and Ki-67 IHC.

The inconsistency in reproducing results using the Ki-67 marker include the selection of hot spots and the determination of the percentage of positive cells among the total number of invasive cancers.
In contrast, the pHH3 marker has a unique feature that confirms the presence of mitosis based on the morphology, like shape and structure. This feature allows for better agreement and accuracy in detecting the presence of specific cells, which may improve the reproducibility of results.
Only moderate inter-rater agreement was observed between pHH3 and Ki-67 IHC, being slightly better when assessed visually. Although both pHH3 and Ki-67 markers are commonly used to evaluate proliferating cells, they target different stages of the cell cycle and are not interchangeable. While Ki-67 is visible in all active cell cycle phases, including mitosis, pHH3 is particularly visible during late G2 and the mitotic phase. Consequently, Ki-67 immunolabeling identifies more positive cells than pHH3 staining.
Strong positive correlation was observed between visual assessment of pHH3 and Ki-67 immunolabeling indices.
Digital evaluation was found to be superior in Ki-67 assessment, compared to the digital evaluation of the pHH3 marker. Dark artifacts and staining intensity posed challenges that were not as prevalent with digital assessment of Ki-67. This may be explained due to the nature of the assessment, since Ki-67 positivity is based on the percentage of positively stained nuclei and pHH3 was scored as the total number of positive cells in 10 HPFs. Furthermore, the Ki-67 Quantifier may be therefore better optimized for Ki-67 assessment compared to pHH3. Staining intensities on slides may vary and digital image analysis might be better suited for the Ki-67 marker. The quality of histological slides might differ for pHH3 and Ki-67 and may impact the accuracy of the digital evaluation.
Thus, only a moderate correlation was found between the digital assessment of pHH3 and Ki-67 IHC.

4.5. Identification of new thresholds

The traditional visual quantification method may have certain advantages over digital quantification, such as the ability of the observer to identify subtle

variations in staining intensity, to evaluate staining patterns, and account for any specific morphological characteristics associated with mitotic cells. Moreover, visual assessment enables the utilization of human expertise and judgment during the analysis, which may consider additional factors that automated digital quantification may overlook.

However, it's important to note that digital quantification has its own advantages, such as providing objective, reproducible and precise measurements, thereby reducing subjectivity, and allowing for potential automation of large datasets.

By conducting both visual and digital quantification cluster analyzes, the study gains a comprehensive understanding of the relationship between pHH3 and Ki-67 staining.

The present study revealed that pHH3 staining highlights a larger number of mitotic cells in comparison to HE-staining, suggesting that the use of the current HE cut-off values to score pHH3-stained slides may lead to an overestimation of the histological grade. To identify potential cut-off values for pHH3 and Ki-67, k-medoid cluster analysis was conducted to group data points into homogeneous clusters, with the optimal number of clusters determined to be 3 for both pHH3 and Ki-67. Accordingly, the study proposes the use of 3 subgroups for pHH3 with minimum to maximum values (such as 1 - 59 vs. 21 - 121 vs. 39 - 226) and 3 subgroups for Ki-67 (such as 5,3 - 35 % vs. 26,4 - 55,7 % vs. 49,7 - 84,8 %) as a reasonable approach for further validation.

4.6. Conclusion

In this study we show that pHH3 counts correlate with mitotic count on HE stains but pHH3 is a more sensitive marker, therefore resulting in a higher number of mitoses than traditional HE slides. While digital pathology exhibited excellent performance in quantifying Ki-67 immunohistochemical staining, our study revealed that accurate quantification of pHH3 using digital methods was not feasible.

Subjectivity in the assessment of histological grade might be reduced by using pHH3, however, further confirmation of suitability and superiority of this biomarker is needed.

4.7. Challenges and future trends

The findings of this study should be considered in the light of some limitations. The relatively small sample size of the cohort and the exploratory nature of the study represent important limitations. Therefore, the prognostic impact of pHH3-assisted mitotic count needs to be validated in larger cohorts as part of the GRK 2543 project. At present, these results are not yet applicable to clinical practice.
Another potential limitation is observer bias, which may arise from researchers' role as observers in collecting data. Our data was selected in a non-random manner, and was aimed to create a balanced dataset, thus introducing selection bias. Additionally, demographic risk factors were not taken into consideration in our analysis. Furthermore, although the software utilized in this study was able to provide an estimation of the total number of cells, number of positive cells, and an index value based on these values, standardized comparisons could not be made without measuring the size of the tissue area that was analyzed. To address this issue, it is necessary to calculate the physical size of each analyzed tissue area by converting the number of pixels to mm^2 using a conversion factor based on the resolution and magnification values employed during the whole slide imaging process.
However, it is important to mention, that this calculation was not possible due to influence from several factors. In many cases the scanned WSI had to be resized, resampled, converted or cut before the software could analyze it. To overcome these circumstances, it would be necessary for the software to be able to process the whole slide images or to carry out the conversion calculations on its own.
As pathology is an image-based discipline, precise analysis, and interpretation of the tissue sections is key for correct diagnosis.
Efforts to overcome some of the challenges seen with traditional pathological analysis have led to AI models, which have evolved into better machine learning techniques, such as deep convolutional neural networks, using biologically inspired networks to represent data. These groundbreaking improvements have greatly aided visual and digital image analysis[124]. Machine learning is a field of artificial intelligence where the algorithm is able to learn and adapt to new data but also to extract information or features beyond human visual perception[125].

AI has the power to analyze large amounts of data quickly and might enable the discovery of novel histopathological features. This might help us to better understand the disease better or promote our ability to predict disease progression and consequently tailor individual treatment plans[125,126].

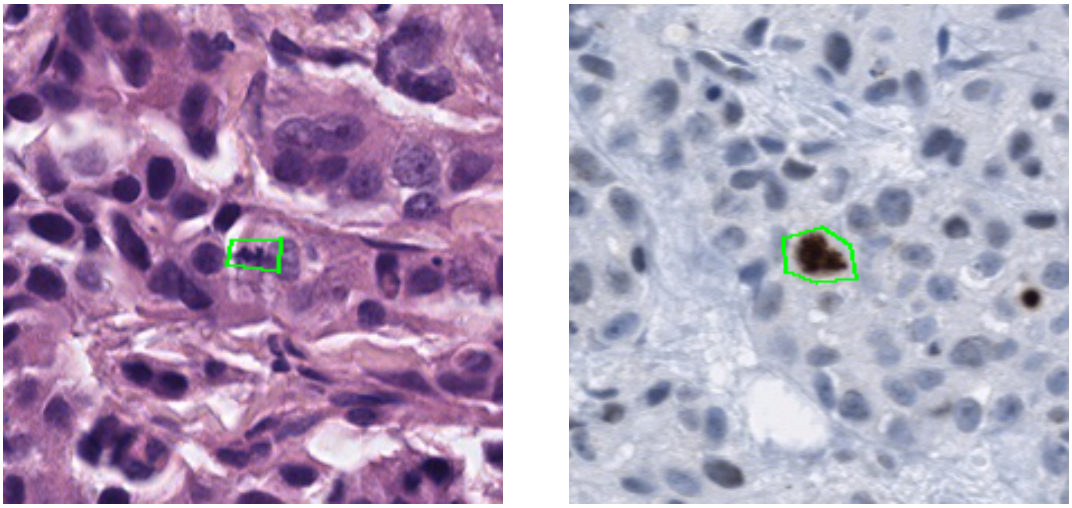
In the future, we aim to explore deep learning algorithms to automatically identify cancerous tissue. Pattern recognition is possible through a training process, for example from datasets annotated by humans.

The dataset collected for this study was prepared as a training set for a deep learning algorithm developed by Simon Holdenried-Krafft and Prof. Dr. Ing. Lensch from the Department of Computer Science in Tübingen. HE stained slides were converted into digital images and cancerous tissue was annotated. The AI algorithm was trained on this dataset for the automated distinction between cancerous and non-cancerous tissue.

Annotated regions were divided into thousands of image patches and fed into a convolutional neural network (CNN).

IHC staining for ER, PR, HER2 and Ki-67 for 74 patients were visually assessed under the microscope and digitized.

Our future aim is to predict breast cancer biomarkers directly from HE slides - potentially bypassing the need for immunohistochemical staining altogether. A similar approach has already been conducted by Couture et al. in 2018, with promising results[127].



a) b)

Figure 16: a) An image patch created of an annotated HE-stained slide. A mitotic figure is circled in green. b) An image patch created of an annotated pHH3 slide, mitotic indices are stained bown. The pHH3-positive cell is circled in green.

5. Summary

Mitotic count is considered to be the best available proliferation marker and the most important grading component to predict prognosis in invasive breast carcinoma. However, it is believed that it is also the cause of subjectivity in grade estimation based on the Elston and Ellis Classification. Thus, quantifiable methods such as immunohistochemistry (pHH3 and Ki-67) based analysis methods hold great potential in supporting histological grading.

In this study, we aimed to evaluate the accuracy and interobserver variability of the Elston and Ellis Score and to compare the conventional method of mitotic count on HE-stained slides with visual and digital assessment of pHH3 and Ki-67 as quantifiable immunohistochemical markers in NST breast carcinoma.

Visual assessment of pHH3 demonstrated a higher sensitivity in detecting mitotic indices than mitotic count on HE-stained slides, suggesting that the visual and digital assessment of mitotic count on HE-stained slides and the evaluation of mitotic indices on pHH3-stained slides are not equivalent approaches. Nevertheless, considering pHH3 immunolabeling, visual assessment allows for better identification of the area with the highest mitotic activity due to the distinct contrast between stained elements.

A strong correlation was observed between the visual and digital approaches, indicating the potential of the digital method as a viable alternative to conventional counting methods. Despite the superior performance of digital pathology in quantifying Ki-67 immunohistochemical staining, our study demonstrated that it was not feasible for accurate quantification of pHH3 due to the software's limited ability to distinguish between dark artifacts and positive cells. Moreover, we propose cut-off values for pHH3 und Ki-67 based on quantitative assessment. The pHH3 marker may potentially improve subjectivity in histological grade and Ki-67 assessment, but additional validation is necessary to establish its suitability as a biomarker.

5.1 Zusammenfassung

Die Mitosenzahl ist eine wichtige Komponente der Elston und Ellis Klassifikation, welche - neben der Kernmorpholgie und der Tubulusdifferenzierung - der histologischen Beurteilung von invasivem Brustkrebs dient. Da die histologische Untersuchung immer noch der Subjektivität der visuellen Untersuchung unterliegt, wurde in dieser Studie die Mitosezahl mit den quantifizierbaren Markern Ki-67 und pHH3, welche die Proliferation der Tumorzellen widerspiegeln, verglichen. Zur histologischen Untersuchung wurden die Gewebeproben unter dem Mikroskop betrachtet und anschließend digitalisiert und durch eine computerbasierte quantitative Bildanalyse evaluiert. Die immunhistochemisch gefärbten Schnitte wurden mit visueller und digitaler Bildanalyse untersucht und verglichen, mit dem Fernziel die Sensitivität und Spezifität der Tumordiagnostik und -klassifizierung durch computerbasierte Algorithmen zu verbessern.

Die visuelle Bewertung von pHH3 zeigte eine höhere Anzahl an Mitosen als auf HE-gefärbten Objektträgern, was darauf zurückzuführen ist, dass pHH3 ein sensitiverer Marken für Mitosen ist. Die Übereinstimmung zwischen der Anzahl der Mitosen und der visuellen Beurteilung von pHH3 und Ki-67 gefärbten Schnitten zeigte sich als mäßig und zwischen der Anzahl der Mitosen und der digitalen Beurteilung von pHH3 und Ki-67 als gering. Visuelle und digitale Methoden korrelierten stark, was auf das Potenzial der digitalen Methode als praktikable Alternative zu den herkömmlichen Zählmethoden hinweist. Trotz der überlegenen Leistung der digitalen Analyse bei der Quantifizierung der immunhistochemischen Ki-67-Färbung zeigte unsere Studie, dass eine genaue Quantifizierung von pHH3 aufgrund der begrenzten Fähigkeit der Software, zwischen dunklen Artefakten und positiven Zellen zu unterscheiden, nicht möglich ist. Darüber hinaus schlagen wir, basierend auf einer Clusteranalyse Cut-off-Werte für pHH3 und Ki67 vor.

Die Anwendung der pHH3 Färbung könnte die Subjektivität der histologischen Klassifikation senken, bedarf in Zukunft jedoch weiterer Untersuchungen.

6. Erklärung zum Eigenanteil

Hiermit versichere ich, Diana Andrea Silimon, dass ich die dem Fachbereich Medizin zur Promotionsprüfung eingereichte Arbeit mit dem Titel „Comparison of visual and digital assessment of mitotic count, Ki-67 and pHH3 immunostaining in breast cancer“ selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht habe. Die statistische Analyse wurde von Lina Maria Serna Higuita vom Universitätsklinikum Tübingen, Institut für Klinische Epidemiologie und angewandte Biometrie und von mir durchgeführt.

Das Projekt wurde von der Ethik-Kommission (Ethik-Kommission an der Medizinisches Fakultät der Eberhard Karls Universität und am Universitätsklinikum Tübingen) beurteilt und am 17.08.2021 genehmigt, die Projekt Nummer lautet 547/2021BO2.

Ich habe bisher an keiner in- oder ausländischen medizinischen Fakultät ein Gesuch um Zulassung zur Promotion eingereicht noch die vorliegende Arbeit als Dissertation vorgelegt.

7. References

1. Elston CW, Ellis IO. 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*. Nov 1991;19(5):403-10. doi:10.1111/j.1365-2559.1991.tb00229.x

2. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer*. Sep 1957;11(3):362. doi:10.1038/bjc.1957.43

3. Metzger-Filho O, Ferreira AR, Jeselsohn R, et al. Mixed Invasive Ductal and Lobular Carcinoma of the Breast: Prognosis and the Importance of Histologic Grade. *Oncologist*. Jul 2019;24(7):e441-e449. doi:10.1634/theoncologist.2018-0363

4. Rakha EA, Reis-Filho JS, Baehner F, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res*. 2010;12(4):207. doi:10.1186/bcr2607

5. Rakha EA, El-Sayed ME, Lee AHS, et al. Prognostic Significance of Nottingham Histologic Grade in Invasive Breast Carcinoma. *Journal of Clinical Oncology*. 2008;26(19):3153-3158. doi:10.1200/jco.2007.15.5986

6. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. Mar 2017;67(2):93-99. doi:10.3322/caac.21388

7. Tan PH, Ellis I, Allison K, et al. The 2019 World Health Organization classification of tumours of the breast. *Histopathology*. 2020;77(2):181-185. doi:10.1111/his.14091

8. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological Image Analysis: A Review. *IEEE Reviews in Biomedical Engineering*. 2009;2:147-171. doi:10.1109/RBME.2009.2034865

9. Meyer JS, Cosatto E, Graf HP. Mitotic index of invasive breast carcinoma. Achieving clinically meaningful precision and evaluating tertial cutoffs. *Arch Pathol Lab Med*. Nov 2009;133(11):1826-33. doi:10.5858/133.11.1826

10. Harvey JM, de Klerk NH, Sterrett GF. Histological Grading in Breast Cancer: Interobserver Agreement, and Relation to Other Prognostic Factors Including Ploidy. *Pathology - Journal of the RCPA*. 1992;24(2):63-68. doi:10.3109/00313029209063625

11. Boiesen P, Bendahl PO, Anagnostaki L, et al. Histologic grading in breast cancer--reproducibility between seven pathologic departments. South Sweden

Breast Cancer Group. *Acta Oncol*. 2000;39(1):41-5. doi:10.1080/028418600430950

12. van Diest PJ, Baak JP, Matze-Cok P, et al. Reproducibility of mitosis counting in 2,469 breast cancer specimens: results from the Multicenter Morphometric Mammary Carcinoma Project. *Hum Pathol*. Jun 1992;23(6):603-7. doi:10.1016/0046-8177(92)90313-r

13. Ibrahim A, Lashen A, Toss M, Mihai R, Rakha E. Assessment of mitotic activity in breast cancer: revisited in the digital pathology era. *Journal of Clinical Pathology*. 2022;75(6):365-372. doi:10.1136/jclinpath-2021-207742

14. Baak JPA, Diest PJv, Voorhorst FJ, et al. Prospective Multicenter Validation of the Independent Prognostic Value of the Mitotic Activity Index in Lymph Node–Negative Breast Cancer Patients Younger Than 55 Years. *Journal of Clinical Oncology*. 2005;23(25):5993-6001. doi:10.1200/jco.2005.05.511

15. Tsuda H, Akiyama F, Kurosumi M, et al. Evaluation of the Interobserver Agreement in the Number of Mitotic Figures Breast Carcinoma as Simulation of Quality Monitoring in the Japan National Surgical Adjuvant Study of Breast Cancer (NSAS-BC) Protocol. *Japanese journal of cancer research*. 2000;91(4):451-457.

16. Simpson JF, Dutt PL, Page DL. Expression of mitoses per thousand cells and cell density in breast carcinomas: A proposal. *Human Pathology*. 1992/06/01/ 1992;23(6):608-611. doi:10.1016/0046-8177(92)90314-S

17. Lee LH, Yang H, Bigras G. Current breast cancer proliferative markers correlate variably based on decoupled duration of cell cycle phases. *Sci Rep*. May 30 2014;4:5122. doi:10.1038/srep05122

18. Tetzlaff MT, Curry JL, Ivan D, et al. Immunodetection of phosphohistone H3 as a surrogate of mitotic figure count and clinical outcome in cutaneous melanoma. *Modern Pathology*. 2013/09/01 2013;26(9):1153-1160. doi:10.1038/modpathol.2013.59

19. Ribalta T, McCutcheon IE, Aldape KD, Bruner JM, Fuller GN. The mitosis-specific antibody anti-phosphohistone-H3 (PHH3) facilitates rapid reliable grading of meningiomas according to WHO 2000 criteria. *Am J Surg Pathol*. Nov 2004;28(11):1532-6. doi:10.1097/01.pas.0000141389.06925.d5

20. Tsuta K, Liu DC, Kalhor N, Wistuba, II, Moran CA. Using the mitosis-specific marker anti-phosphohistone H3 to assess mitosis in pulmonary neuroendocrine carcinomas. *Am J Clin Pathol*. Aug 2011;136(2):252-9. doi:10.1309/ajcpdxfopxgef0rp

21. Colman H, Giannini C, Huang L, et al. Assessment and prognostic significance of mitotic index using the mitosis marker phospho-histone H3 in low

and intermediate-grade infiltrating astrocytomas. *Am J Surg Pathol*. May 2006;30(5):657-64. doi:10.1097/01.pas.0000202048.28203.25

22. Skaland I, Janssen EAM, Gudlaugsson E, et al. 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*. 2007/12/01 2007;20(12):1307-1315. doi:10.1038/modpathol.3800972

23. Baak JP, Gudlaugsson E, Skaland I, et al. Proliferation is the strongest prognosticator in node-negative breast cancer: significance, error sources, alternatives and comparison with molecular prognostic markers. *Breast Cancer Res Treat*. May 2009;115(2):241-54. doi:10.1007/s10549-008-0126-y

24. Dessauvagie BF, Thomas C, Robinson C, Frost FA, Harvey J, Sterrett GF. Validation of mitosis counting by automated phosphohistone H3 (PHH3) digital image analysis in a breast carcinoma tissue microarray. *Pathology*. 2015/06/01/ 2015;47(4):329-334. doi:10.1097/PAT.0000000000000248

25. Trihia H, Murray S, Price K, et al. Ki-67 expression in breast carcinoma: its association with grading systems, clinical parameters, and other prognostic factors--a surrogate marker? *Cancer*. Mar 1 2003;97(5):1321-31. doi:10.1002/cncr.11188

26. Viale G, Regan MM, Mastropasqua MG, et al. Predictive value of tumor Ki-67 expression in two randomized trials of adjuvant chemoendocrine therapy for node-negative breast cancer. *J Natl Cancer Inst*. Feb 6 2008;100(3):207-12. doi:10.1093/jnci/djm289

27. Jensen EV, and Jacobson, H.I. Basic guides to the mechanism of estrogen action. . *Recent Prog Horm Res*. 1962;(18):387-414.

28. Luporsi E, André F, Spyratos F, et al. Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review. *Breast Cancer Res Treat*. Apr 2012;132(3):895-915. doi:10.1007/s10549-011-1837-z

29. Nielsen TO, Leung SCY, Rimm DL, et al. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *J Natl Cancer Inst*. Jul 1 2021;113(7):808-819. doi:10.1093/jnci/djaa201

30. Thomssen C, Balic M, Harbeck N, Gnant M. St. Gallen/Vienna 2021: A Brief Summary of the Consensus Discussion on Customizing Therapies for Women with Early Breast Cancer. *Breast Care*. 2021;16(2):135-143. doi:10.1159/000516114

31. Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst*. Nov 16 2011;103(22):1656-64. doi:10.1093/jnci/djr393

32. Puri M, Hoover SB, Hewitt SM, et al. Automated Computational Detection, Quantitation, and Mapping of Mitosis in Whole-Slide Images for Clinically Actionable Surgical Pathology Decision Support. *J Pathol Inform*. 2019;10:4. doi:10.4103/jpi.jpi_59_18

33. Robertson S, Azizpour H, Smith K, Hartman J. Digital image analysis in breast pathology-from image processing techniques to artificial intelligence. *Transl Res*. Apr 2018;194:19-35. doi:10.1016/j.trsl.2017.10.010

34. Klauschen F, Wienert S, Schmitt WD, et al. Standardized Ki67 Diagnostics Using Automated Scoring—Clinical Validation in the GeparTrio Breast Cancer Study. *Clinical Cancer Research*. 2015;21(16):3651-3657. doi:10.1158/1078-0432.Ccr-14-1283

35. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. May 2021;71(3):209-249. doi:10.3322/caac.21660

36. Katalinic A, Eisemann N, Kraywinkel K, Noftz MR, Hübner J. Breast cancer incidence and mortality before and after implementation of the German mammography screening program. *Int J Cancer*. Aug 1 2020;147(3):709-718. doi:10.1002/ijc.32767

37. Anttinen J, Kautiainen H, Kuopio T. Role of mammography screening as a predictor of survival in postmenopausal breast cancer patients. *British Journal of Cancer*. 2006/01/01 2006;94(1):147-151. doi:10.1038/sj.bjc.6602895

38. ELLIS IO, GALEA M, BROUGHTON N, LOCKER A, BLAMEY RW, ELSTON CW. Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up. *Histopathology*. 1992;20(6):479-489. doi:10.1111/j.1365-2559.1992.tb01032.x

39. Zhu H, Doğan BE. American Joint Committee on Cancer's Staging System for Breast Cancer, Eighth Edition: Summary for Clinicians. *Eur J Breast Health*. Jul 2021;17(3):234-238. doi:10.4274/ejbh.galenos.2021.2021-4-3

40. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*. 2000/08/01 2000;406(6797):747-752. doi:10.1038/35021093

41. Rakha EA, Ellis IO. Modern classification of breast cancer: should we stick with morphology or convert to molecular profile characteristics. *Adv Anat Pathol*. Jul 2011;18(4):255-67. doi:10.1097/PAP.0b013e318220f5d1

42. Vuong D, Simpson PT, Green B, Cummings MC, Lakhani SR. Molecular classification of breast cancer. *Virchows Archiv*. 2014/07/01 2014;465(1):1-14. doi:10.1007/s00428-014-1593-7

43. Sinn HP, Kreipe H. A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition. *Breast Care (Basel)*. May 2013;8(2):149-54. doi:10.1159/000350774

44. Schwartz AM, Henson DE, Chen D, Rajamarthandan S. Histologic grade remains a prognostic factor for breast cancer regardless of the number of positive lymph nodes and tumor size: a study of 161 708 cases of breast cancer from the SEER Program. *Arch Pathol Lab Med*. Aug 2014;138(8):1048-52. doi:10.5858/arpa.2013-0435-OA

45. Leo McGuire W. Current status of estrogen receptors in human breast cancer. *Cancer*. 1975;36(S2):638-644.

46. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. Jan 9 1987;235(4785):177-82. doi:10.1126/science.3798106

47. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *Jama*. 2006;295(21):2492-2502.

48. Spitale A, Mazzola P, Soldini D, Mazzucchelli L, Bordoni A. Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland. *Annals of Oncology*. 2009/04/01/ 2009;20(4):628-635. doi:10.1093/annonc/mdn675

49. Duffy MJ, Harbeck N, Nap M, et al. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). *Eur J Cancer*. Apr 2017;75:284-298. doi:10.1016/j.ejca.2017.01.017

50. Lumachi F, Santeufemia DA, Basso SM. Current medical treatment of estrogen receptor-positive breast cancer. *World J Biol Chem*. Aug 26 2015;6(3):231-9. doi:10.4331/wjbc.v6.i3.231

51. Johnston SJ, Cheung KL. Endocrine Therapy for Breast Cancer: A Model of Hormonal Manipulation. *Oncol Ther*. Dec 2018;6(2):141-156. doi:10.1007/s40487-018-0062-x

52. Figueroa-Magalhães MC, Jelovac D, Connolly R, Wolff AC. Treatment of HER2-positive breast cancer. *Breast*. Apr 2014;23(2):128-136. doi:10.1016/j.breast.2013.11.011

53. Veta M, van Diest PJ, Jiwa M, Al-Janabi S, Pluim JPW. Mitosis Counting in Breast Cancer: Object-Level Interobserver Agreement and Comparison to an

Automatic Method. *PLOS ONE*. 2016;11(8):e0161286. doi:10.1371/journal.pone.0161286

54. Gerdes J, Li L, Schlueter C, et al. Immunobiochemical and molecular biologic characterization of the cell proliferation-associated nuclear antigen that is defined by monoclonal antibody Ki-67. *Am J Pathol*. Apr 1991;138(4):867-73.

55. Soliman NA, Yussif SM. Ki-67 as a prognostic marker according to breast cancer molecular subtype. *Cancer Biol Med*. Dec 2016;13(4):496-504. doi:10.20892/j.issn.2095-3941.2016.0066

56. Lindboe CF, Torp SH. Comparison of Ki-67 equivalent antibodies. *J Clin Pathol*. Jun 2002;55(6):467-71. doi:10.1136/jcp.55.6.467

57. de Azambuja E, Cardoso F, de Castro G, Jr., et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. *Br J Cancer*. May 21 2007;96(10):1504-13. doi:10.1038/sj.bjc.6603756

58. Cheang MC, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst*. May 20 2009;101(10):736-50. doi:10.1093/jnci/djp082

59. Viale G, Giobbie-Hurder A, Regan MM, et al. Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endocrine-responsive breast cancer: results from Breast International Group Trial 1-98 comparing adjuvant tamoxifen with letrozole. *J Clin Oncol*. Dec 1 2008;26(34):5569-75. doi:10.1200/jco.2008.17.0829

60. Cuzick J, Dowsett M, Pineda S, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *J Clin Oncol*. Nov 10 2011;29(32):4273-8. doi:10.1200/jco.2010.31.2835

61. Yeo B, Zabaglo L, Hills M, Dodson A, Smith I, Dowsett M. Clinical utility of the IHC4+C score in oestrogen receptor-positive early breast cancer: a prospective decision impact study. *British Journal of Cancer*. 2015/07/01 2015;113(3):390-395. doi:10.1038/bjc.2015.222

62. Dong Z, Bode AM. The role of histone H3 phosphorylation (Ser10 and Ser28) in cell growth and cell transformation. *Mol Carcinog*. Jun 2006;45(6):416-21. doi:10.1002/mc.20220

63. Kim JY, Jeong HS, Chung T, et al. The value of phosphohistone H3 as a proliferation marker for evaluating invasive breast cancers: A comparative study with Ki67. *Oncotarget*. Sep 12 2017;8(39):65064-65076. doi:10.18632/oncotarget.17775

64. Voss SM, Riley MP, Lokhandwala PM, Wang M, Yang Z. Mitotic Count by Phosphohistone H3 Immunohistochemical Staining Predicts Survival and Improves Interobserver Reproducibility in Well-differentiated Neuroendocrine Tumors of the Pancreas. *The American Journal of Surgical Pathology*. 2015;39(1):13-24. doi:10.1097/pas.0000000000000341

65. Mirzaiian E, Tabatabaei Ghods ZS, Tavangar SM, et al. Utility of PHH3 in Evaluation of Mitotic Index in Breast Carcinoma and Impact on Tumor Grade. *Asian Pac J Cancer Prev*. Jan 1 2020;21(1):63-66. doi:10.31557/apjcp.2020.21.1.63

66. Cui X, Harada S, Shen D, Siegal GP, Wei S. The Utility of Phosphohistone H3 in Breast Cancer Grading. *Applied Immunohistochemistry & Molecular Morphology*. 2015;23(10):689-695. doi:10.1097/pai.0000000000000137

67. Sørlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences*. 2001;98(19):10869-10874. doi:doi:10.1073/pnas.191367098

68. Kondov B, Milenkovikj Z, Kondov G, et al. Presentation of the Molecular Subtypes of Breast Cancer Detected By Immunohistochemistry in Surgically Treated Patients. *Open Access Macedonian Journal of Medical Sciences*. 06/06 2018;6(6):961-967. doi:10.3889/oamjms.2018.231

69. Gnant M, Harbeck N, Thomssen C. St. Gallen 2011: Summary of the Consensus Discussion. *Breast Care (Basel)*. 2011;6(2):136-141. doi:10.1159/000328054

70. Prat A, Pineda E, Adamo B, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast*. Nov 2015;24 Suppl 2:S26-35. doi:10.1016/j.breast.2015.07.008

71. Senkus E, Kyriakides S, Ohno S, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. Sep 2015;26 Suppl 5:v8-30. doi:10.1093/annonc/mdv298

72. Yersal O, Barutca S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J Clin Oncol*. Aug 10 2014;5(3):412-24. doi:10.5306/wjco.v5.i3.412

73. von Minckwitz G, Schneeweiss A, Loibl S, et al. Neoadjuvant carboplatin in patients with triple-negative and HER2-positive early breast cancer (GeparSixto; GBG 66): a randomised phase 2 trial. *Lancet Oncol*. Jun 2014;15(7):747-56. doi:10.1016/s1470-2045(14)70160-3

74. Schneeweiss A, Denkert C, Fasching PA, et al. Diagnosis and Therapy of Triple-Negative Breast Cancer (TNBC) - Recommendations for Daily Routine

Practice. *Geburtshilfe Frauenheilkd*. Jun 2019;79(6):605-617. doi:10.1055/a-0887-0285

75. Gecer B, Aksoy S, Mercan E, Shapiro LG, Weaver DL, Elmore JG. Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern Recognit*. Dec 2018;84:345-356. doi:10.1016/j.patcog.2018.07.022

76. Ehteshami Bejnordi B, Mullooly M, Pfeiffer RM, et al. Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Mod Pathol*. Oct 2018;31(10):1502-1512. doi:10.1038/s41379-018-0073-z

77. Cruz-Roa A, Gilmore H, Basavanhally A, et al. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. *Sci Rep*. Apr 18 2017;7:46450. doi:10.1038/srep46450

78. Araújo T, Aresta G, Castro E, et al. Classification of breast cancer histology images using Convolutional Neural Networks. *PLoS One*. 2017;12(6):e0177544. doi:10.1371/journal.pone.0177544

79. Barsha NA, Rahman A, Mahdy MRC. Automated detection and grading of Invasive Ductal Carcinoma breast cancer using ensemble of deep learning models. *Comput Biol Med*. Dec 2021;139:104931. doi:10.1016/j.compbiomed.2021.104931

80. Chan RC, To CKC, Cheng KCT, Yoshikazu T, Yan LLA, Tse GM. Artificial intelligence in breast cancer histopathology. *Histopathology*. 2023;82(1):198-210. doi:10.1111/his.14820

81. Irshad H. Automated mitosis detection in histopathology using morphological and multi-channel statistics features. *J Pathol Inform*. 2013;4:10. doi:10.4103/2153-3539.112695

82. Mathew T, Ajith B, Kini JR, Rajan J. Deep learning-based automated mitosis detection in histopathology images for breast cancer grading. *International Journal of Imaging Systems and Technology*. 2022;32(4):1192-1208. doi:10.1002/ima.22703

83. Sohail A, Khan A, Wahab N, Zameer A, Khan S. A multi-phase deep CNN based mitosis detection framework for breast cancer histopathological images. *Scientific Reports*. 2021/03/18 2021;11(1):6215. doi:10.1038/s41598-021-85652-1

84. Badve SS. Artificial intelligence in breast pathology – dawn of a new era. *npj Breast Cancer*. 2023/01/31 2023;9(1):5. doi:10.1038/s41523-023-00507-4

85. Patel A, Balis UGJ, Cheng J, et al. Contemporary Whole Slide Imaging Devices and Their Applications within the Modern Pathology Department: A

Selected Hardware Review. *J Pathol Inform*. 2021;12:50. doi:10.4103/jpi.jpi_66_21

86. Ying X, Monticello TM. Modern Imaging Technologies in Toxicologic Pathology: An Overview. *Toxicologic Pathology*. 2006/12/01 2006;34(7):815-826. doi:10.1080/01926230600918983

87. Treanor D. Virtual slides: an introduction. *Diagnostic Histopathology*. 2009/02/01/ 2009;15(2):99-103. doi:10.1016/j.mpdhp.2009.01.006

88. Remmele W, Stegner HE. [Recommendation for uniform definition of an immunoreactive score (IRS) for immunohistochemical estrogen receptor detection (ER-ICA) in breast cancer tissue]. *Pathologe*. May 1987;8(3):138-40. Vorschlag zur einheitlichen Definition eines Immunreaktiven Score (IRS) für den immunhistochemischen Ostrogenrezeptor-Nachweis (ER-ICA) im Mammakarzinomgewebe.

89. Kassambara A. *Practical guide to cluster analysis in R: Unsupervised machine learning*. vol 1. Sthda; 2017.

90. Genestie C, Zafrani B, Asselain B, et al. Comparison of the prognostic value of Scarff-Bloom-Richardson and Nottingham histological grades in a series of 825 cases of breast cancer: major importance of the mitotic count as a component of both grading systems. *Anticancer Res*. Jan-Feb 1998;18(1b):571-6.

91. van Diest PJ, van der Wall E, Baak JPA. Prognostic value of proliferation in invasive breast cancer: a review. *Journal of Clinical Pathology*. 2004;57(7):675. doi:10.1136/jcp.2003.010777

92. Medri L, Volpi A, Nanni O, et al. Prognostic Relevance of Mitotic Activity in Patients with Node-Negative Breast Cancer. *Modern Pathology*. 2003/11/01 2003;16(11):1067-1075. doi:10.1097/01.MP.0000093625.20366.9D

93. Frierson HF, Jr., Wolber RA, Berean KW, et al. Interobserver reproducibility of the Nottingham modification of the Bloom and Richardson histologic grading scheme for infiltrating ductal carcinoma. *Am J Clin Pathol*. Feb 1995;103(2):195-8. doi:10.1093/ajcp/103.2.195

94. Meyer JS, Alvarez C, Milikowski C, et al. Breast carcinoma malignancy grading by Bloom–Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index. *Modern Pathology*. 2005/08/01 2005;18(8):1067-1078. doi:10.1038/modpathol.3800388

95. Harbeck N, Thomssen C, Gnant M. St. Gallen 2013: brief preliminary summary of the consensus discussion. *Breast Care (Basel)*. May 2013;8(2):102-9. doi:10.1159/000351193

96. Sikka M, Agarwal S, Bhatia A. Interobserver agreement of the Nottingham histologic grading scheme for infiltrating duct carcinoma breast. *Indian J Cancer*. Jun-Dec 1999;36(2-4):149-53.

97. Robbins P, Pinder S, de Klerk N, et al. Histological grading of breast carcinomas: a study of interobserver agreement. *Hum Pathol*. Aug 1995;26(8):873-9. doi:10.1016/0046-8177(95)90010-1

98. Ginter PS, Idress R, D'Alfonso TM, et al. Histologic grading of breast carcinoma: a multi-institution study of interobserver variation using virtual microscopy. *Modern Pathology*. 2021/04/01 2021;34(4):701-709. doi:10.1038/s41379-020-00698-2

99. Rabe K, Snir OL, Bossuyt V, Harigopal M, Celli R, Reisenbichler ES. Interobserver variability in breast carcinoma grading results in prognostic stage differences. *Human Pathology*. 2019/12/01/ 2019;94:51-57. doi:10.1016/j.humpath.2019.09.006

100. Zhang R, Chen HJ, Wei B, et al. Reproducibility of the Nottingham modification of the Scarff-Bloom-Richardson histological grading system and the complementary value of Ki-67 to this system. *Chin Med J (Engl)*. Aug 5 2010;123(15):1976-82.

101. Bueno-de-Mesquita JM, Nuyten DS, Wesseling J, van Tinteren H, Linn SC, van de Vijver MJ. The impact of inter-observer variation in pathological assessment of node-negative breast cancer on clinical risk assessment and patient selection for adjuvant systemic treatment. *Ann Oncol*. Jan 2010;21(1):40-7. doi:10.1093/annonc/mdp273

102. Ladstein RG, Bachmann IM, Straume O, Akslen LA. Ki-67 expression is superior to mitotic count and novel proliferation markers PHH3, MCM4 and mitosin as a prognostic factor in thick cutaneous melanoma. *BMC Cancer*. Apr 14 2010;10:140. doi:10.1186/1471-2407-10-140

103. Liu N, Song S-Y, Jiang J-B, Wang T-J, Yan C-X. The prognostic role of Ki-67/MIB-1 in meningioma: A systematic review with meta-analysis. *Medicine*. 2020;99(9):e18644. doi:10.1097/md.0000000000018644

104. Urruticoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. *J Clin Oncol*. Oct 1 2005;23(28):7212-20. doi:10.1200/jco.2005.07.501

105. Spyratos F, Ferrero-Poüs M, Trassard M, et al. Correlation between MIB-1 and other proliferation markers: clinical implications of the MIB-1 cutoff value. *Cancer*. Apr 15 2002;94(8):2151-9. doi:10.1002/cncr.10458

106. Denkert C, Loibl S, Müller BM, et al. Ki67 levels as predictive and prognostic parameters in pretherapeutic breast cancer core biopsies: a

translational investigation in the neoadjuvant GeparTrio trial. *Ann Oncol*. Nov 2013;24(11):2786-93. doi:10.1093/annonc/mdt350

107. Madani SH, Payandeh M, Sadeghi M, Motamed H, Sadeghi E. The correlation between Ki-67 with other prognostic factors in breast cancer: A study in Iranian patients. *Indian J Med Paediatr Oncol*. Apr-Jun 2016;37(2):95-9. doi:10.4103/0971-5851.180136

108. Mohammed AA. Quantitative assessment of Ki67 expression in correlation with various breast cancer characteristics and survival rate; cross sectional study. *Annals of Medicine and Surgery*. 2019/12/01/ 2019;48:129-134. doi:10.1016/j.amsu.2019.11.005

109. Hashmi AA, Hashmi KA, Irfan M, et al. Ki67 index in intrinsic breast cancer subtypes and its association with prognostic parameters. *BMC Research Notes*. 2019/09/23 2019;12(1):605. doi:10.1186/s13104-019-4653-x

110. Yousif M, van Diest PJ, Laurinavicius A, et al. Artificial intelligence applied to breast pathology. *Virchows Archiv*. 2022/01/01 2022;480(1):191-209. doi:10.1007/s00428-021-03213-3

111. Alataki A, Zabaglo L, Tovey H, Dodson A, Dowsett M. A simple digital image analysis system for automated Ki67 assessment in primary breast cancer. *Histopathology*. 2021;79(2):200-209. doi:10.1111/his.14355

112. del Rosario Taco Sanchez M, Soler-Monsó T, Petit A, et al. Digital quantification of KI-67 in breast cancer. *Virchows Archiv*. 2019/02/01 2019;474(2):169-176. doi:10.1007/s00428-018-2481-3

113. Pons L, Hernández-León L, Altaleb A, et al. Conventional and digital Ki67 evaluation and their correlation with molecular prognosis and morphological parameters in luminal breast cancer. *Scientific Reports*. 2022/05/17 2022;12(1):8176. doi:10.1038/s41598-022-11411-5

114. Goto H, Tomono Y, Ajiro K, et al. Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. *J Biol Chem*. Sep 3 1999;274(36):25543-9. doi:10.1074/jbc.274.36.25543

115. Wei Y, Mizzen CA, Cook RG, Gorovsky MA, Allis CD. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena. *Proc Natl Acad Sci U S A*. Jun 23 1998;95(13):7480-4. doi:10.1073/pnas.95.13.7480

116. Bossard C, Jarry A, Colombeix C, et al. Phosphohistone H3 labelling for histoprognostic grading of breast adenocarcinomas and computer-assisted determination of mitotic index. *J Clin Pathol*. Jul 2006;59(7):706-10. doi:10.1136/jcp.2005.030452

117. Duregon E, Molinaro L, Volante M, et al. Comparative diagnostic and prognostic performances of the hematoxylin-eosin and phospho-histone H3

mitotic count and Ki-67 index in adrenocortical carcinoma. *Modern Pathology*. 2014/09/01 2014;27(9):1246-1254. doi:10.1038/modpathol.2013.230

118. Thareja S, Zager JS, Sadhwani D, et al. Analysis of tumor mitotic rate in thin metastatic melanomas compared with thin melanomas without metastasis using both the hematoxylin and eosin and anti-phosphohistone 3 IHC stain. *Am J Dermatopathol*. Jan 2014;36(1):64-7. doi:10.1097/DAD.0b013e31829433b6

119. van Steenhoven JEC, Kuijer A, Kornegoor R, et al. Assessment of tumour proliferation by use of the mitotic activity index, and Ki67 and phosphohistone H3 expression, in early-stage luminal breast cancer. *Histopathology*. Oct 2020;77(4):579-587. doi:10.1111/his.14185

120. Kim YJ, Ketter R, Steudel WI, Feiden W. Prognostic significance of the mitotic index using the mitosis marker anti-phosphohistone H3 in meningiomas. *Am J Clin Pathol*. Jul 2007;128(1):118-25. doi:10.1309/hxunag34b3cefdu8

121. Caicedo JC, Goodman A, Karhohs KW, et al. Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nature Methods*. 2019/12/01 2019;16(12):1247-1253. doi:10.1038/s41592-019-0612-7

122. van Dooijeweert C, van Diest PJ, Ellis IO. Grading of invasive breast carcinoma: the way forward. *Virchows Archiv*. 2022/01/01 2022;480(1):33-43. doi:10.1007/s00428-021-03141-2

123. Ghaznavi F, Evans A, Madabhushi A, Feldman M. Digital imaging in pathology: whole-slide imaging and beyond. *Annu Rev Pathol*. Jan 24 2013;8:331-59. doi:10.1146/annurev-pathol-011811-120902

124. Zhu J, Liu M, Li X. Progress on deep learning in digital pathology of breast cancer: a narrative review. *Gland Surgery*. 2022;11(4):751-766.

125. Aeffner F, Zarella MD, Buchbinder N, et al. Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. *Journal of Pathology Informatics*. 2019/01/01/ 2019;10(1):9. doi:10.4103/jpi.jpi_82_18

126. Barsoum I, Tawedrous E, Faragalla H, Yousef GM. Histo-genomics: digital pathology at the forefront of precision medicine. *Diagnosis (Berl)*. Aug 27 2019;6(3):203-212. doi:10.1515/dx-2018-0064

127. Couture HD, Williams LA, Geradts J, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer*. 2018;4:30. doi:10.1038/s41523-018-0079-1

8. Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) – Project Number 409474577. RTG 2543: Intraoperative Multisensory Tissue Differentiation in Oncology.

I would like to express my deep appreciation to Professor Dr. Falko Fend for the opportunity to work on this project and for his valuable guidance. His trust, his constant support and help enabled me to develop professionally and personally, for which I am very grateful.

I would like to extend my sincere thanks to my supervisor Ivonne Montes for her support and feedback, without which this work would not have been possible.

Furthermore, I would like to thank Lina Maria Serna Higuita for her help in the analysis of the clinical data.

I would like to express my sincere gratitude to the entire pathology department for the pleasant work atmosphere.

My heartfelt thanks go to the patients who agreed to participate in the present study.