Virchows Archiv
https://doi.org/10.1007/s00428-026-04500-7

ORIGINAL ARTICLE

Optimization of Ki67 digital image analysis in breast cancer by automated tumor area identification

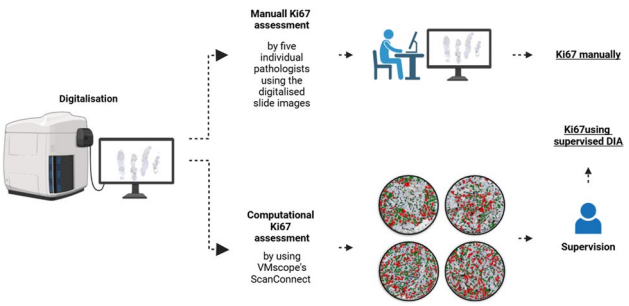
Paul Jank[1,2] · Maximilian J. Krämer[2] · Michael Untch[3] · Mattea Reinisch[4] · Theresa Link[5] · Albert Grass[2] · Jens-Uwe Blohmer[1] · Claus Hanusch[6] · Anne-Sophie Litmeyer[2] · Jens Huober[7] · Dominik Gerber[8] · Christine Solbach[9] · Vesna Bjelic-Radisic[10,11] · Kai Saeger[12] · Andreas Schneeweiss[13] · Kerstin Rhiem[14] · Moritz Gleitsmann[2] · Susanne von Gerlach[2] · Bärbel Felder[15] · Moritz Jesinghaus[2] · Carsten Denkert[2] · Sibylle Loibl[15] · Wolfgang D. Schmitt[16]

Received: 16 October 2025 / Revised: 23 February 2026 / Accepted: 22 March 2026
© The Author(s) 2026, modified publication 2026

Abstract

Accurate assessment of Ki67, a marker of cellular proliferation, is critical for breast cancer diagnostics and treatment decision-making. This study evaluates an automated Ki67-area identification approach, combined with digital image analysis (DIA) for exact Ki67 quantification. A total of 61 tissue samples from breast cancer patients from two clinical trials (GeparSixto and GeparSepto) were analyzed. The supervised DIA workflow employed automated Ki67-stained tumor area identification, followed by automated Ki67 scoring and was quality-controlled by trained pathologists. Comparisons with manual assessments were conducted to evaluate concordance and scoring precision. The DIA approach identified 232 tumor areas across 61 whole slide images (WSIs). The supervised system demonstrated a high correlation with manual scores ($r = 0.78$), and improved precision (standard deviation between evaluated tumor areas) was noted in therapy-naïve samples ($p < 0.001$). While manual assessments predominantly employed stepwise increments (5%), DIA provided finer, more continuous scoring, optimally reflecting the continuous nature of tumor proliferation. Overall, DIA scoring was more robust, with reduced inter-ROI variability and improved reproducibility compared with manual assessment. In conclusion, the study highlights the potential of supervised DIA to enhance Ki67 scoring accuracy and standardization. The methodology's effectiveness in addressing standardized Ki67 scoring suggests its utility in clinical diagnostic workflows. This approach contributes towards improved integration of computational pathology into routine practice for improved breast cancer management.

Graphical Abstract



Keywords Computational pathology · Ki67 · Breast cancer

Paul Jank and Maximilian J. Krämer shared first authorship and contributed equally.

Extended author information available on the last page of the article

Published online: 23 April 2026

Introduction

Sustaining uncontrolled proliferation is one of the hallmarks of cancer influencing tumor progression and patient prognosis [10]. In breast cancer, Ki67, a nucleic marker expressed during late S-phase, G2 and M-phase, is widely used to calculate the proliferation index. Precise assessment of proliferation plays a pivotal role in pathological routine diagnostics [9]. A comparison of clinical guidelines such as those from the European Society for Medical Oncology (ESMO), the German S3 Guideline, and the St. Gallen Consensus Conference reveals no global consensus regarding factors such as the number of tumor cells or tumor areas to assess or Ki67 cutoffs for differentiating between low- and high-proliferating tumors [2, 21], Leitlinienprogramm [16].

Besides its well-known prognostic impact, Ki67 has been described as a predictive biomarker for therapy response in the neoadjuvant setting [6–8]. Especially in hormone receptor (HR)-positive, human epidermal growth factor receptor 2 (HER2)-negative early breast cancer, Ki67 is a driving factor for adjuvant chemotherapy indication. Since 2022, Ki67 is approved by the U.S. Food & Drug Administration (FDA) as companion diagnostics, guiding adjuvant therapy with Abemaciclib. Patients with HR-positive, HER2-negative early or metastatic breast cancer showing ≥20% Ki67-positive tumor cells are eligible for adding Abemaciclib to adjuvant standard-of-care endocrine therapy (ET) [11].

Since Ki67 was introduced as both a predictive and therapy-guiding biomarker, as well as a companion diagnostic for therapeutic decision-making, an objective and standardized procedure for assessing Ki67 is essential. According to the German S3 Guidelines, pathologists should assess more than three tumor areas and count approximately 100 to 300 tumor cells, which may require a considerable amount of time. The international Ki67 Working Group (IKWG) published guidelines for the implementation of Ki67 in clinical breast cancer assessment, highlighting the necessity for a standardized approach to Ki67 assessment [19].

Digital software tools for assessing a wide range of immunohistochemical markers are currently under development, aiming to standardize diagnostic procedures that are especially crucial in routine diagnostics. These systems are discussed as offering a significant opportunity to overcome the challenges currently faced in traditional diagnostics [23]. However, despite their potential, many of these systems are difficult to implement in a clinical routine due to complex workflows or poor correlation with traditional scoring methods [3].

The Cognition Master Professional Suite (CMPS) software, a comprehensive toolset developed for clinical use in pathological research and diagnostics, introduced an easily applicable workflow for standardized digital Ki67 scoring that has been validated in different clinical studies [1, 12, 26, 27].

Here, we benchmark a novel supervised digital image analysis (DIA) protocol employing the CMPS using automated identification of representative Ki67-stained regions of interest (ROI) on whole-slide images (WSI), followed by automated Ki67 scoring using the Ki67 Quantifier Clinical module and subsequent pathologist supervision. The primary objective of this study is to compare this WSI-based DIA approach with manual Ki67 assessment performed according to established clinical practice guidelines, focusing on concordance, variability (standard deviations), and time required for each assessment method.

Therefore, we used FFPE tumor samples of two multicenter randomized clinical German Breast Group (GBG) trials, with tissue sections CNB taken prior to and during neoadjuvant chemotherapy (NACT) and surgical resection specimens obtained after completion of NACT.

Materials and methods

Sample collection

Sixty-one randomly selected longitudinal tissue samples were derived from 29 female breast cancer patients enrolled in the clinical GBG trials GeparSixto (G6) and GeparSepto (G7) [22, 24]. Where applicable, we assessed longitudinal samples with up to three time points from one patient: therapy-naïve sample from core needle biopsies (CNB) prior to neoadjuvant treatment (pretherapeutic), mid-treatment (intratherapeutic) samples using CNB after 2 cycles of chemotherapy (GeparSixto) or from window phase (one dose of anti-HER2 treatment in GeparSepto) and surgical resection specimens obtained after completion of neoadjuvant chemotherapy (postNACT, postherapeutic) according to the clinical trial.

All patients provided written informed consent for study participation and biomaterial collection for translational research. Ethical approvals for translational research projects were given by Charité University Hospital Berlin (EA1/139/05), as well as Philipps-University Marburg (38/20). All clinical trials underlying this study were conducted in accordance with the Declaration of Helsinki.

Immunohistochemistry and digitalization

Immunohistochemical staining targeting Ki67 was conducted on whole-slide sections of 2 µm using monoclonal mouse anti-human-Ki67 antibody (Clone MIB-1, Dako Denmark A/S) with a 1:100 dilution and EDTA antigen

retrieval. Enhancing visualization, hematoxylin (SS3301, Dako, Denmark A/S) was employed as counterstain on a Leica Bond Max (Dako, Denmark A/S) automatized stainer. Subsequently, tissue slide digitalization was carried out on a M8 slide scanner (PreciPoint GmbH, Germany) at a magnification of 20× generating whole-slide images (WSI).

Manual Ki67 assessment

Manual Ki67 scoring was performed on the digitized WSI on a digital screen. Five pathologists were instructed to score according to their standard practice in accordance with St. Gallen Consensus guidelines (2021) as well as the German S3 guideline (version 4.4, 2021) [4], Leitlinienprogramm [16]. According to these recommendations, pathologists assessed at least three representative high-power fields and counted a minimum of 100 tumor cells per field, covering areas with representative staining intensities and proportions of positive cells. Malignant cells were counted as positive if they showed low to strong nuclear staining. Individual Ki67 values were calculated as the number of positively stained tumor cells in relation to the total number of assessed tumor cells. For each slide, the manual Ki67 score used as a gold standard for comparative data analysis was based on the mean scores of all pathologists.

Supervised digital image analysis

Supervised Ki67 scoring followed a three-step protocol using two modules of the Cognition Master Professional Suite (VMscope GmbH, Berlin, Germany), and a manual review in-between. First, Ki67 ROI detection with the software module "Scan Connect" (VMscope GmbH, Berlin, Germany) was used to automatically identify up to four single tumor areas per WSI. Each area was standardized to 600 × 600 dpi. The automatic selection of valid ROI regions was based on the staining intensity of the immunohistochemistry (IHC) signal after color deconvolution. To optimize the region selection, the candidate regions were prioritized by proportion of the positive IHC signal, and also their cell–cell as well as possible ROI–ROI proximity to each other, in order to better account for the tumor heterogeneity. These unsupervised images were quality controlled by a trained pathologist, who removed remaining regions with staining artifacts and ROI with poor tumor tissue quality. The final fully automated step of the Ki67 cell detection and scoring was performed by the Ki-67 Quantifier clinical module (VMscope GmbH, Berlin, Germany) for each individual tumor area. The module's algorithmic workflow can be summarized as follows: First, colour deconvolution is performed to separate the Ki67 signal from the counterstain. Subsequently, the colour space of both signal components is analyzed to determine image-specific threshold values. This is followed by contour-based segmentation and contour length analysis to define image-specific object sizes. The detected contours are then classified into tumor cells and other objects (e.g., lymphocytes) based on size, color, and shape features. In a subsequent step, tumor cells are classified as positive or negative according to the previously defined signal components. Finally, all relevant objects are counted, the Ki67 ratio is calculated, and the result is displayed.

Data comparison and statistical analysis

Concordance analyses were performed using Ki67 values obtained from both manual assessment and digital image analysis (DIA). Within the manual Ki67 scoring approach, statistical comparisons were conducted using the Ki67 values derived from five independent pathologists. For the DIA approach, Ki67 values were calculated based on automatically identified regions of interest (ROIs). Statistical analyses were conducted using GraphPad Prism 6 (GraphPad Software Inc, Boston, USA) software and R (version 4.1.2). Paired sample two-sided t-tests were performed comparing manual and supervised DIA scores. The assumption of variance homogeneity was assessed using F-tests. Sample correlation was calculated by Pearson's method, assigned brackets indicate the 95% confidence interval. A significance level of $p < 0.05$ was used to determine statistical significance. Scoring accuracies are a measure of the standard deviation (precision), assessing inter-observer (manual Ki67 assessment) and inter-ROI (un-/supervised DIA) deviations.

To evaluate clinical coherence, WSIs were grouped into Ki67 groups: low (< 10%), intermediate (≤ 25%), and high (> 25%), based on German S3 Guidelines from 2020 for breast cancer [15] as well as Ki67 groups according to the St. Gallen consensus: low (< 5%), intermediate (5–30%), and high (> 30%) [4], Leitlinienprogramm [16]. Concordances were calculated as the ratio of samples classified identically using both approaches to the total number of samples within each group, as classified manually.

Results

Flow of samples and basic characteristics

Within 61 whole slide images (WSIs) Ki67 ROI detection identified 232 individual tumor areas, with three to four ROIs per WSI (four ROIs: 55/61 (90.2%), three ROIs: 2/61 (3.3%), two ROIs: 2/61 (3.3%) and one ROI: 2/61 (3.3%) (Fig. 1). Subsequent Ki67 scoring using DIA was performed on an average of 1124 tumor cells per sample. ROI identification, Ki67 scoring and manual review took around 1.5 min per WSI. The manual review section took up most of the

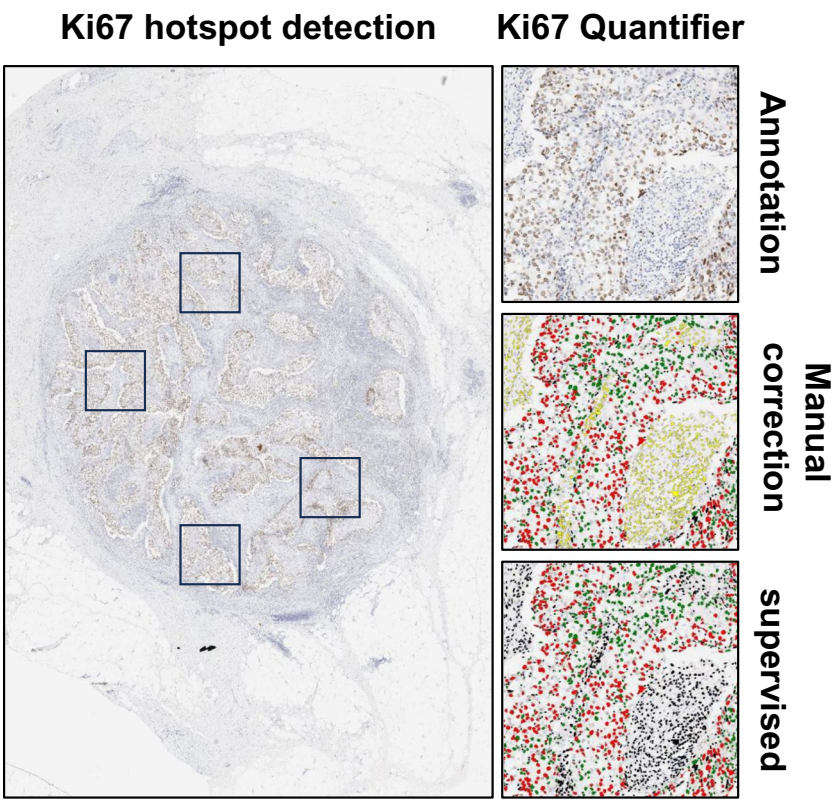


Fig. 1 Illustration of the supervised DIA workflow: Up to four hotspots are selected from the WSI. Within these hotspot regions, the Ki67 Quantifier distinguishes the Ki67 hotspot detection, tumor cells, and non-tumor cells, further classifying Ki67-positive tumor cells (red) and Ki67-negative tumor cells (green). Areas with misclassifications can be refined through manual annotation (yellow), with excluded zones subsequently marked in black

active assessment time since steps upfront are performed in an automated manner. In a direct comparison, the reduction in times becomes apparent since pathologists reported an active Ki67 assessment time of between five and ten minutes.

During the supervision of the DIA by an experienced pathologist, 44 (18.9%) of 232 tumor areas were partially or entirely excluded due to staining artifacts, large necrotic areas, or other quality-reducing factors. This led to exclusion of eight WSIs on which no tumor areas remained (7 intratherapeutic and 1 post-NACT sample). These samples were excluded from further analysis, resulting in a data set of 53 WSIs. The supervision process by trained pathologists significantly reduced the Ki67 inter-ROI deviation from unsupervised 9.83 to 6.11 in the supervised setting ($p=0.008$). Precision variance comparison showed a significantly reduced variance following supervision (unsupervised 97.4, supervised 43.2 F-test $p=0.003$) (Supp. Figure 1).

Comparison of manual and supervised-DIA Ki67 assessment

Analysis of the mean Ki67 scores revealed a notable difference in scoring methodology (Fig. 2). Manual scoring predominantly produced scores that were assigned in increments of approximately 5% intervals, reflecting the semi-quantitative assessment that is typically performed by human evaluators. In contrast, the supervised DIA system delivers stepless values.

Our results show that manual and supervised DIA Ki67 scorings are comparable according to their mean Ki67 values per sample with no significant differences (t-test $p=0.333$; Pearson's correlation: $r=0.781$, 95% CI 0.648–0.868) (Fig. 3A, B). However, computational assessment exhibited lower mean standard deviation (precision) per sample, with a statistically significant lower mean per-sample standard deviation (t-test $p=0.011$), whereas in precision variance no significant difference was observed (manual 38.7, supervised DIA 43.2 F-test $p=0.682$) (Fig. 3C). Upon stratifying tissue samples based on timepoint of treatment, both cohorts displayed notable consistency between pathologists and computational systems: pre-NACT (t-test $p=0.338$; Pearson's correlation: $r=0.707$, 95% CI 0.460–0.853) and intratherapeutic (t-test $p=0.455$; Pearson's correlation: $r=0.820$, 95% CI 0.609–0.923) (Fig. 3D–H). Supervised DIA scoring demonstrated superior scoring accuracy compared with the manual approach for pre-NACT samples. Mean scoring accuracy for supervised DIA was significantly lower (two-sided t-test $p<0.001$), as well as precision variance (manual 24.1, supervised DIA 10.4 F-test $p=0.029$) (Fig. 3F). The intratherapeutic cohort showed no significant differences in mean accuracy (t-test $p=0.732$) while precision variance was significantly better for manual assessment (manual 30.7, supervised DIA 90.8 F-test $p=0.008$) (Fig. 3I).

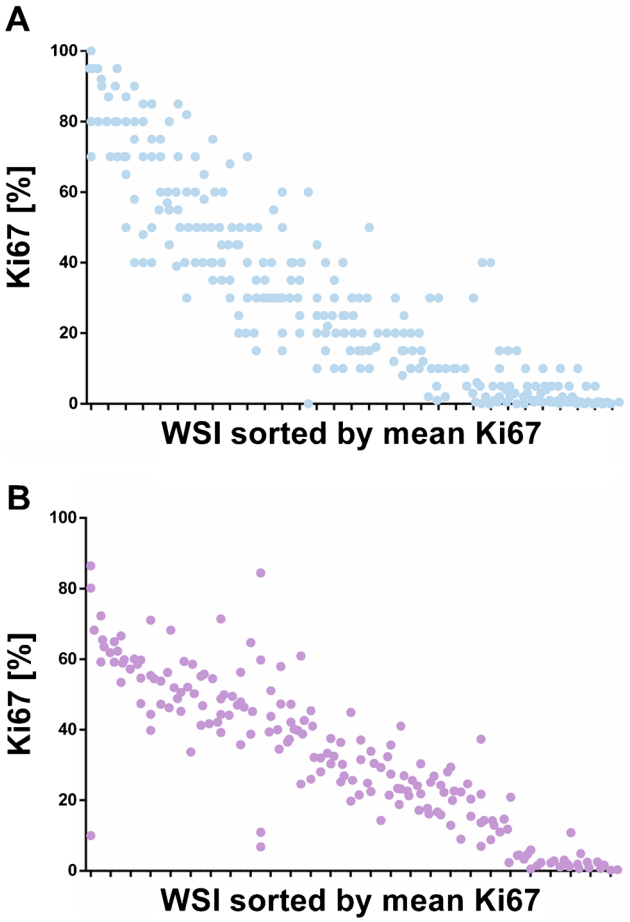


Fig. 2 Comparison of mean Ki67 scores using manual Ki67 assessment and supervised DIA. Each dot represents a whole slide images Ki67 score assessed either manually by a pathologist (**A**, blue) or by supervised DIA (**B**, purple). Samples are not matched, but sorted by mean Ki67 values indicating the sensitivity in scoring

Precision assessment and clinical coherence in Ki67 subgroups

Based on the national German S3 guideline for breast cancer, Ki67 values were categorized into three groups: Ki67 low (< 10%), intermediate (≤ 25%), and high (> 25%). Manual scoring identified 12/53 (22.6%) WSIs as Ki67 low, 14/53 (26.4%) as Ki67 intermediate, and 27/53 (50.9%) as Ki67 high. Supervised DIA scoring identified 11/53 (20.8%) WSIs as Ki67 low, 10/53 (18.9%) as Ki67 intermediate, and 32/53 (60.4%) as Ki67 high. As indicated in Fig. 4, assessment of scoring precision showed low inter-observer and inter-ROI area deviations for Ki67 low and intermediate, respectively (man. and semi-aut.: low sd = 4.5 and 1.7; int. sd = 6.7 and 4.6; high sd = 12.5 and 8.1).

Classification based on the St. Gallen consensus (low (< 5%), intermediate (5–30%), and high (> 30%)) shifted more samples into the intermediate group (Fig. 4B). Manual scoring identified 7/53 (13.2%) WSIs as Ki67 low, 21/53 (39.6%) as Ki67 intermediate, and 25/53 (47.2%) as Ki67 high. Supervised DIA scoring identified 10/53 (18.9%) WSIs as Ki67 low, 17/53 (32.1%) as Ki67 intermediate and 26/53 (49.1%) as Ki67 high. We observed a coherent trend in inter-observer and inter-ROI area deviations compared to the German S3 consensus (man. and DIA: low sd = 3.4 and 1.1; int. sd = 7.9 and 5.7; high sd = 11.9 and 8.3).

The clinical scoring coherence between the two methods was 71.7% (38/53 cases) (Supp. Table 1). However, Ki67 group stratification revealed differences in scoring concordance between groups: Ki67 low (CR = 66.7% (8/12 cases)), intermediate (CR = 35.7% (5/14 cases)), and high (CR = 92.6% (25/27 cases)).

Discussion

Digital software tools for the assessment of a broad range of immunohistochemical markers are currently under development with the aim of standardizing diagnostic procedures in routine clinical practice. These systems are discussed as offering a significant opportunity to overcome the challenges currently faced in traditional diagnostics [23]. Ki67, in particular, lacks standardized scoring procedures, highlighting the potential benefits of digital tools. In this study, we evaluated a supervised DIA approach that combines automated Ki67 ROI identification with image analysis, followed by professional supervision.

Our results show almost comparable Ki67 scores between manual assessment and supervised DIA, with lower inter-ROI deviation (supervised DIA) compared to inter-observer variances. To our knowledge, this is the first time that supervised DIA was employed using automated Ki67-stained ROI detection on whole-slide images with subsequent supervised Ki67 scoring using longitudinal breast cancer tumor material. Previous studies relied on manual identification of representative tissue areas (ROI) and annotation, followed by DIA [3, 13, 20, 25]. In a study conducted by Alataki et al., the CMPS was employed using the Ki67 Quantifier module (VMscope, Berlin, Germany) on both core biopsies and excision specimens. The tumor areas analyzed were selected manually, in contrast to the approach employed in this study. Alataki et al. reported a correlation of $r = 0.92$ CI (0.87–0.94) for core biopsies and $r = 0.95$ CI (0.86–0.98) for excision specimens [1].

Although a high correlation between manual Ki67 assessment and DIA was observed, the process of manual tissue area selection requires considerable effort and expertise. In contrast, automated approaches offer potential for increased efficiency and time savings. Moreover, manual selection may be influenced by various factors, including tissue and staining quality, as well as necrosis or therapy-related tissue

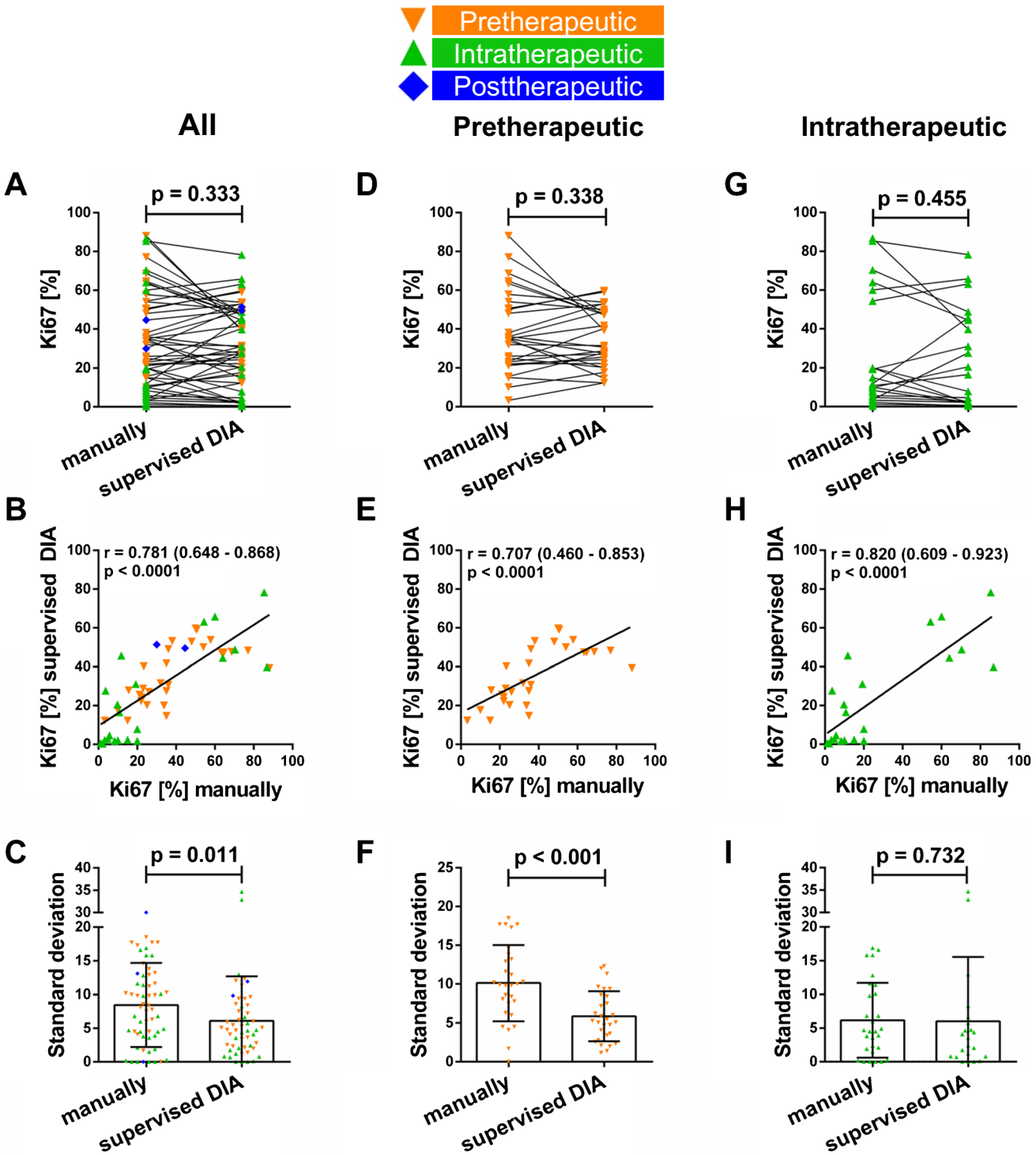


Fig. 3 Comparison of Ki67 scoring performed manually or using supervised DIA A,B,C mean Ki67 differences between manual vs. supervised DIA assessment. **A** All samples, **B** pretherapeutic (orange), **C** intratherapeutic (green) tissue. **D**, **E**, **F** Pearson's correlation between manual and supervised DIA, brackets indicate the 95% confidence interval. **D** All samples, **E** pretherapeutic (orange), and **F** intratherapeutic (green) tissue. **G**, **H**, **I** Intra-tumor area and interrater standard deviation. **G** All samples, **H** pretherapeutic (orange), **I** intratherapeutic (green) tissue. All *p*-values were calculated by two-sided paired *t*-testing

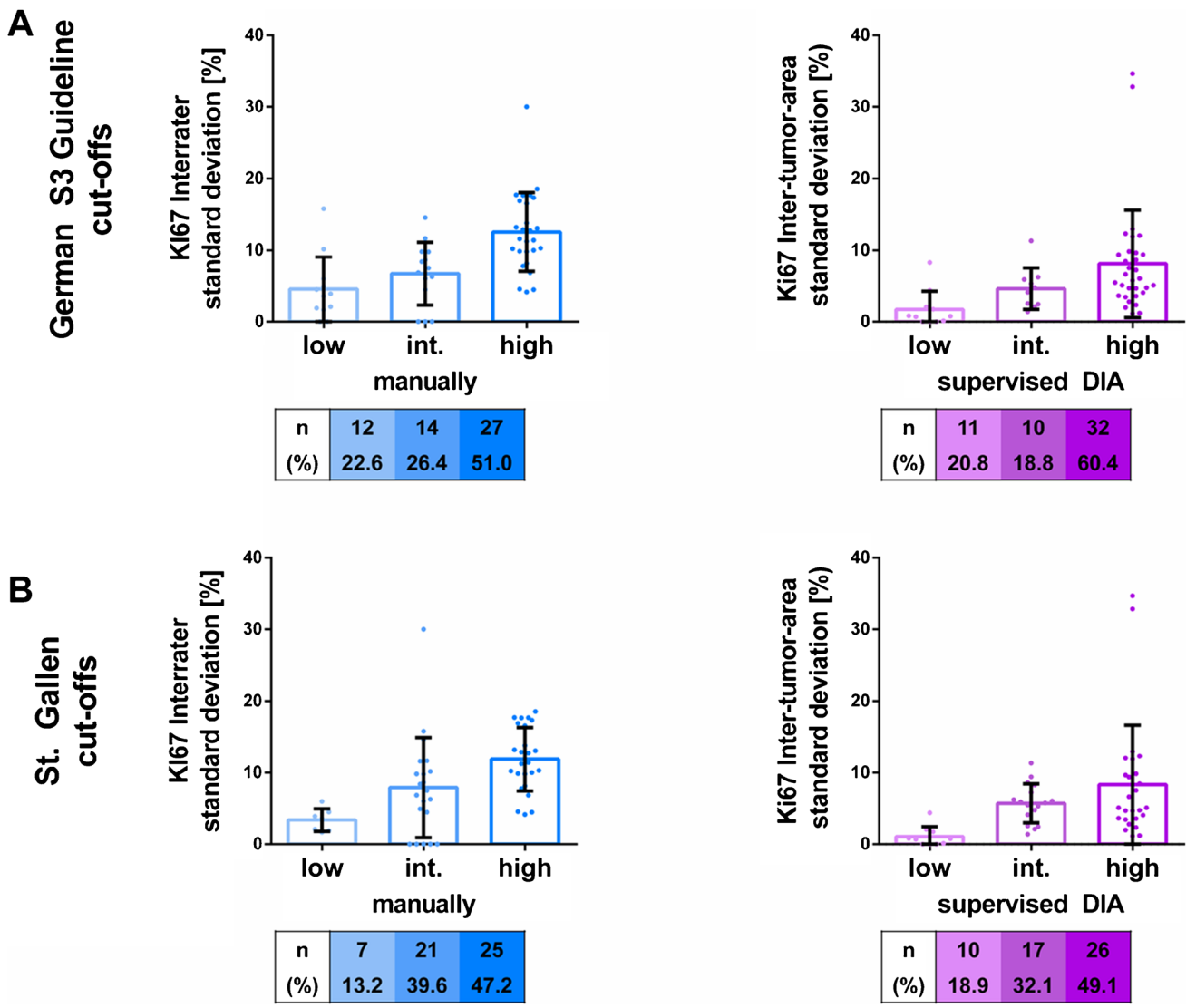


Fig. 4 Interrater (manual assessment) and inter-tumor area standard deviations (supervised DIA) in Ki67 classification groups: **A** following German S3 Guideline cutoffs low > 10%, intermediate (int.) ≤ 25% and high < 25%; **B** following St. Gallen cutoffs: low > 5%, intermediate (int.) 5–30% and high > 30%

remodeling (e.g., fibrosis). The use of a system that automatically identifies ROIs has the potential to reduce the manual workload associated with this process, while also improving the methods' independence from subjective annotations, as discussed by Dawe et al. [5].

To address the issue of standardization in Ki67 assessment, we employed supervised DIA. In the initial stage, Ki67 ROIs were identified autonomously using the Scan Connect software module, and subsequently, the Ki67 Quantifier autonomously scored Ki67 on the selected tumor areas. Finally, the data were reviewed by a trained pathologist.

Our results demonstrated that unsupervised DIA basically matches with the manual approach but significantly improved with pathologist supervision (= supervised DIA). In comparing manual and supervised DIA scoring for Ki67, most pathologists approximated scores in 5% increments. Given that no standard Ki67 analysis was proposed, minor discrepancies were to be expected, in line with the findings of previous studies [18]. Comparatively, supervised DIA provided more continuous scoring, promoting standardization over approximation.

When interrater and inter-ROI standard deviations were compared to assess method precision, supervised DIA demonstrated higher scoring homogeneity regarding Ki67 scores. Stratification into pre-NACT and intratherapeutic tissues revealed high correlations in both subgroups, with significantly higher precision of pre-NACT scores compared to the precision of intratherapeutic scores. A recent

study using DIA with manual area determination on post-treatment specimens has reported challenges in scoring, due to tissue artifacts and lowered numbers of cells [1]. Chemotherapy induces a reduction in proliferative cells and an increase in reactive morphologic changes, which may cause accurate scoring more challenging. Nevertheless, the less significant results in intra-therapeutic tumor samples may be due to a lower case number. DIA are mostly trained on therapy-naïve tumor tissue with high cellularity, leading to a training bias. Due to the reduced cell content and irregular shape of the cells after chemotherapy, this data does not align with the training data, which consequently reduces data quality.

In clinical routine, Ki67 scores are employed for the purposes of case subtyping, e.g. differentiating luminal A and B hormone receptor subtypes and for making therapeutic decision-making. Consequently, the clinical validity of the supervised DIA approach must be addressed. Ki67 index categorization is heterogenous across different working groups. According to German S3 Guideline for breast cancer, Ki67 is classified into three index groups: low (< 10%), intermediate (≤ 25%), and high (> 25%). In clinically sensitive groups with low (< 10%) and intermediate (≤ 25%) (German S3 guideline) Ki67 values, minimal interrater and inter-ROI deviations were observed, indicating high accuracy for Ki67 scoring in these clinically relevant groups. The degree of concordance between manual and supervised DIA classification of Ki67 varied across the three classification categories. The highest concordance was observed in cases with high Ki67 levels (93%), followed by those with low Ki67 levels (67%). The lowest concordance was observed for Ki67-intermediate (Ki67-int) cases. It is already well known in routine clinical practice that scoring Ki67, particularly in the case of Ki67-int, is a challenging task [14, 17]. These findings were reproduced when Ki67 groups are classified according to the St. Gallen consensus.

The limited size of the study cohort may reduce the generalizability of our results, particularly in the context of clinical applications. Further validation studies are needed to ensure the robustness and reliability.

In conclusion, our results demonstrate that the supervised DIA method, which employs autonomous ROI detection, automated Ki67 scoring, and manual supervision, achieves high concordance with traditional clinical scoring. Supervised DIA shows lower variability compared to manual scoring; therefore, DIA should be preferred, especially for clinically relevant decision-making.

Abbreviations CMPS: Cognition Master Professional Suite; DIA: Digital image analysis; ET: Endocrine therapy; FDA: U.S. Food & Drug Administration; G6: GeparSixto; G7: GeparSepto; GBG: German Breast Group; HER2: Human epidermal growth factor receptor 2; HR: Hormone receptor; ICH: Immunohistochemistry; IKWG: International Ki67 Working Group; Ki67-int: Ki67-intermediate; ROI: Representative tissue areas; SD: Standard deviation; WSI: Whole slide images

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00428-026-04500-7.

Acknowledgements We would like to thank all patients, physicians, scientists, and all other employees participating in the trials and the translational project. A special thanks goes to Sofia Born and Aran Talaei for intensive help in data curation and figure production. We are grateful to Marion Kalden, Mareike Meier, and Viktoria Wischmann for excellent technical assistance.

Author contribution Conceptualization: CD and PJ. Data curation: MJK, PJ, AG, AL, MG, SG, and AR. Formal analysis: MJK and PJ. Funding acquisition: CD. Investigation: PJ, JL, DT, JB, CD, and SL. Methodology: PJ and MJK. Supervision: CD, PJ, and WDS. Writing original draft: MJK and PJ. Review and editing: PJ, MJK, BVS, AG, AL, MU, DG, AS, KS, MG, JF, SG, BF, MR, SL, CD, and WDS. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This project has been supported by the Cogno-Scan grant (grant ID: 13GW0207A) from the German Federal Ministry of Education and Research.

Data Availability All relevant data are within the paper and its Supporting Information files. The data underlying the results presented in the study are available from German Breast Group, some restriction apply due to confidentiality of patient data. Since these data are derived from a prospective clinical trial with ongoing follow up there are legal and ethical restrictions to share sensitive patient related data publicly. Interested groups may use the "Cooperation Proposal Form" on https://www.gbg.de/en/research/trafo.php. Data can be requested in context of a translational research project by sending the form to trafo@gbg.de. Translational research proposals are approved by the GBG scientific board.

Declarations

Conflict of interest P. Jank reports grants and travel expenses from Gilead Sciences GmbH outside the submitted work. C. Denkert reports other support from MSD Oncology, personal fees from Daiichi Sankyo and AstraZeneca, and grants from Myriad Genetics and German Breast Group outside the submitted work. C. Denkert also reports a patent for VMscope digital pathology software with royalties paid, a patent for WO2020109570A1 issued, a patent for WO2015114146A1 issued, and a patent for WO2010076322A1 issued. W. D. Schmitt received speakers' honoraria from AstraZeneca, Bundesverband Deutscher Pathologen, Deutsche Akademie für Senologie, GlaxoSmithKline, NOGGO, and Roche outside of this study and adboard honoraria from AstraZeneca and Myriad outside of this study, and financial funding from Myriad outside of this study. K. Saeger is CEO of VMscope GmbH, which is the producer and distributor of the CMPS software and its modules "Ki67 Quantifier clinical" and "Scan Connect." J. Huober received honoraria from Lilly, Novartis, Roche, Pfizer, AstraZeneca, Gilead, Daiichi, Stemline, MSD, and AbbVie and reports consulting or advisory relationships with Lilly, Novartis, Roche, Pfizer, AstraZeneca, Gilead, Daiichi, and AbbVie and received travel expenses from Roche, Daiichi, Gilead, and AstraZeneca. M. Reinisch reports honoraria and/or consultancy fees from AstraZeneca, Daiichi Sankyo, Gilead, Lilly, MSD, Novartis, Pfizer, Roche, OnkowissenTV, Seagen, Streamed Up, and Somatex. M. Reinisch also reports travel support from AstraZeneca, Daiichi Sankyo, Gilead, Lilly, Novartis, Pfizer, and Roche. T. Link reports Honoraria for Lectures or presentations from Amgen, Roche,

MSD, Novartis, Pfizer, Lilly, GSK, Gilead, Astra Zeneca, Daiichi Sankyo, Stemline, Seagen; Advisory Board/Advise from MSD, Roche, Pfizer, Lilly, Myriad, Esai, GSK, Gilead, Daiichi Sankyo, Roche, Astra Zeneca and support for attending meetings and/or travel from Pfizer, Astra Zeneca, Gilead, Daiichi Sankyo, Stemline, Seagen, MSD. C. Hanusch reports Speakers Bureau from Astra Zeneca, Novartis, Roche. S. declares to be GBG Forschungs GmbH employee (CEO); The institution receives grants and/or contracts from the following entities: AbbVie, AICME, AstraZeneca, Agendia, Amgen, Bayer, BeiGene, Bicycle, BioNTech, Celgene/BMS, Cellcuity, Chugai Pharmaceuticals, Corcept Therapeutics, DSI, Exact Science, Gilead, GSK, Incyte, Lilly, Medscape, MICE ideas, Molecular Health, Jiangsu Hengrui, MSD, Novartis, Olema, Peerview, Pierre Fabre, Pfizer, Relay, Roche, Sanofi, Seagen, Stemline/Menarini, JAZZ Pharmaceuticals; receives support for attending meetings and/or travel from DSI, Gilead, Roche, Lilly, ESMO, SGBCC, ASCO, AGO Kommission Mamma; GBG Forschungs GmbH has following /patents planned, issued or pending: EP14153692.0, EP21152186.9, EP18209672 and EP24210258 and receives licensing fees from VM Scope GmbH. B. Felder declares to be GBG Forschungs GmbH employee; The institution receives grants and/or contracts from the following entities: AbbVie, AICME, AstraZeneca, Agendia, Amgen, Bayer, BeiGene, Bicycle, BioNTech, Celgene/BMS, Cellcuity, Chugai Pharmaceuticals, Corcept Therapeutics, DSI, Exact Science, Gilead, GSK, Incyte, Lilly, Medscape, MICE ideas, Molecular Health, Jiangsu Hengrui, MSD, Novartis, Olema, Peerview, Pierre Fabre, Pfizer, Relay, Roche, Sanofi, Seagen, Stemline/Menarini, JAZZ Pharmaceuticals; receives support for attending meetings and/or travel from DSI, Gilead, Roche, Lilly, ESMO, SGBCC, ASCO, AGO Kommission Mamma; GBG Forschungs GmbH has following /patents planned, issued or pending: EP14153692.0, EP21152186.9, EP18209672 and EP24210258 and receives licensing fees from VM Scope GmbH.

Declaration of generative AI and AI-assisted technologies in the writing process During the preparation of this work, the authors used ChatGPT OpenAI (OpenAI Ireland Ltd., Dublin, Ireland) for spell and grammar checking. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Alataki A, Zabaglo L, Tovey H, Dodson A, Dowsett M (2021) A simple digital image analysis system for automated Ki67 assessment in primary breast cancer. Histopathology 79:200–209
2. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, Zackrisson S, Senkus E, Esmo Guidelines Committee (2019) Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-updagger'. Ann Oncol 30:1194–220
3. Catteau X, Zindy E, Bouri S, Noel JC, Salmon I, Decaestecker C (2023) Comparison between manual and automated assessment of Ki-67 in breast carcinoma: test of a simple method in daily practice. Technol Cancer Res Treat 22:15330338231169604
4. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, Thurlimann B, Senn HJ, Members Panel (2015) Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. Ann Oncol 26:1533–1546
5. Dawe M, Shi W, Liu TY, Lajkosz K, Shibahara Y, Gopal NEK, Geread R, Mirjahanmardi S, Wei CX, Butt S, Abdalla M, Manolescu S, Liang SB, Chadwick D, Roehrl MHA, McKee TD, Adeoye A, McCready D, Khademi A, Liu FF, Fyles A, Done SJ (2024) Reliability and variability of Ki-67 digital image analysis methods for clinical diagnostics in breast cancer. Lab Invest 104:100341
6. Denkert C, Budczies J, von Minckwitz G, Wienert S, Loibl S, Klauschen F (2015) Strategies for developing Ki67 as a useful biomarker in breast cancer. Breast 24(Suppl 2):S67-72
7. Denkert C, von Minckwitz G (2014) Reply to Ki67 in breast cancer: a useful prognostic marker! Ann Oncol 25:542–543
8. Dowsett M, Nielsen TO, Rimm DL, Hayes DF, Group International Ki67 in Breast Cancer Working (2022) Ki67 as a companion diagnostic: good or bad news? J Clin Oncol 40:3796–3799
9. Gerdes J, Becker MH, Key G, Cattoretti G (1992) Immunohistological detection of tumour growth fraction (Ki-67 antigen) in formalin-fixed and routinely processed tissues. J Pathol 168:85–86
10. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144:646–674
11. Johnston SRD, Harbeck N, Hegg R, Toi M, Martin M, Shao ZM, Zhang QY, Martinez Rodriguez JL, Campone M, Hamilton E, Sohn J, Guarneri V, Okada M, Boyle F, Neven P, Cortes J, Huober J, Wardley A, Tolaney SM, Cicin I, Smith IC, Frenzel M, Headley D, Wei R, San Antonio B, Hulstijn M, Cox J, O'Shaughnessy J, Rastogi P, E. Committee Members monarch, Investigators (2020) Abemaciclib Combined With Endocrine Therapy for the Adjuvant Treatment of HR+, HER2-, Node-Positive, High-Risk, Early Breast Cancer (monarchE). J Clin Oncol 38:3987–98
12. Klauschen F, Wienert S, Schmitt WD, Loibl S, Gerber B, Blohmer JU, Huober J, Rudiger T, Erbstosser E, Mehta K, Lederer B, Dietel M, Denkert C, von Minckwitz G (2015) Standardized Ki67 diagnostics using automated scoring - clinical validation in the GeparTrio breast cancer study. Clin Cancer Res 21:3651–3657
13. Koopman T, Buikema HJ, Hollema H, de Bock GH, van der Vegt B (2018) Digital image analysis of Ki67 proliferation index in breast cancer using virtual dual staining on whole tissue sections: clinical validation and inter-platform agreement. Breast Cancer Res Treat 169:33–42
14. Kwon AY, Park HY, Hyeon J, Nam SJ, Kim SW, Lee JE, Yu JH, Lee SK, Cho SY, Cho EY (2019) Practical approaches to automated digital image analysis of Ki-67 labeling index in 997 breast carcinomas and causes of discordance with visual assessment. PLoS ONE 14:e0212309
15. Leitlinienprogramm-Onkologie (2020) Interdisziplinäre S3-Leitlinie für die Früherkennung, Diagnostik, Therapie und Nachsorge des Mammakarzinoms. AWMF online. https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Mammakarzinom_4_0/Version_4.3/LL_Mammakarzinom_Langversion_4.3.pdf. Accessed 14 Aug 2021
16. Leitlinienprogramm Onkologie (2021) S3-Leitlinie Mammakarzinom | Version 4.4 | Juni 2021. Leitlinienprogramm Onkologie. https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Mammakarzinom_4_0/Version_4.4/LL_Mammakarzinom_Langversion_4.4.pdf. Accessed 14 Aug 2021
17. Leung SCY, Nielsen TO, Zabaglo L, Arun I, Badve SS, Bane AL, Bartlett JMS, Borgquist S, Chang MC, Dodson A, Enos RA, Fineberg S, Focke CM, Gao D, Gown AM, Grabau D, Gutierrez C, Hugh JC, Kos Z, Laenkholm AV, Lin MG, Mastropasqua MG, Moriya T, Nofech-Mozes S, Osborne CK, Penault-Llorca FM,

Piper T, Sakatani T, Salgado R, Starczynski J, Viale G, Hayes DF, McShane LM, Dowsett M (2016) Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration. NPJ Breast Cancer 2:16014
18. Li W, Lu N, Chen C, Lu X (2023) Identifying the optimal cutoff point of Ki-67 in breast cancer: a single-center experience. J Int Med Res 51:3000605231195468
19. Nielsen TO, Leung SCY, Rimm DL, Dodson A, Acs B, Badve S, Denkert C, Ellis MJ, Fineberg S, Flowers M, Kreipe HH, Laenkholm AV, Pan H, Penault-Llorca FM, Polley MY, Salgado R, Smith IE, Sugie T, Bartlett JMS, McShane LM, Dowsett M, Hayes DF (2021) Assessment of Ki67 in breast cancer: updated recommendations from the International Ki67 in Breast Cancer Working Group. J Natl Cancer Inst 113:808–819
20. Skjervold AH, Pettersen HS, Valla M, Opdahl S, Bofin AM (2022) Visual and digital assessment of Ki-67 in breast cancer tissue - a comparison of methods. Diagn Pathol 17:45
21. Thomssen C, Balic M, Harbeck N, Gnant M (2021) St. Gallen/Vienna 2021: a brief summary of the consensus discussion on customizing therapies for women with early breast cancer. Breast Care (Basel) 16:135–143
22. Untch M, Jackisch C, Schneeweiss A, Conrad B, Aktas B, Denkert C, Eidtmann H, Wiebringhaus H, Kummel S, Hilfrich J, Warm M, Paepke S, Just M, Hanusch C, Hackmann J, Blohmer JU, Clemens M, Darb-Esfahani S, Schmitt WD, Costa SD, Gerber B, Engels K, Nekljudova V, Loibl S, von Minckwitz G, Group German Breast, Investigators Arbeitsgemeinschaft Gynakologische Onkologie-Breast (2016) Nab-paclitaxel versus solvent-based paclitaxel in neoadjuvant chemotherapy for early breast cancer (GeparSepto-GBG 69): a randomised, phase 3 trial. Lancet Oncol 17:345–356
23. Verghese G, Lennerz JK, Ruta D, Ng W, Thavaraj S, Siziopikou KP, Naidoo T, Rane S, Salgado R, Pinder SE, Grigoriadis A (2023) Computational pathology in cancer diagnosis, prognosis, and prediction - present day and prospects. J Pathol 260:551–563
24. von Minckwitz G, Schneeweiss A, Loibl S, Salat C, Denkert C, Rezai M, Blohmer JU, Jackisch C, Paepke S, Gerber B, Zahm DM, Kummel S, Eidtmann H, Klare P, Huober J, Costa S, Tesch H, Hanusch C, Hilfrich J, Khandan F, Fasching PA, Sinn BV, Engels K, Mehta K, Nekljudova V, Untch M (2014) Neoadjuvant carboplatin in patients with triple-negative and HER2-positive early breast cancer (GeparSixto; GBG 66): a randomised phase 2 trial. Lancet Oncol 15:747–756
25. Wang YX, Wang YY, Yang CG, Bu H, Yang WT, Wang L, Xu WM, Zhao XL, Zhao WX, Li L, Song SL, Yang JL (2020) An interobserver reproducibility analysis of size-set semiautomatic counting for Ki67 assessment in breast cancer. Breast 49:225–232
26. Wienert S, Heim D, Kotani M et al (2013) CognitionMaster: an object-based image analysis framework. Diagn Pathol 8:34. https://link.springer.com/article/10.1186/1746-1596-8-34
27. Wienert S, Heim D, Saeger K, Stenzinger A, Beil M, Hufnagl P, Dietel M, Denkert C, Klauschen F (2012) Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. Sci Rep 2:503

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Paul Jank[1,2] · Maximilian J. Krämer[2] · Michael Untch[3] · Mattea Reinisch[4] · Theresa Link[5] · Albert Grass[2] · Jens-Uwe Blohmer[1] · Claus Hanusch[6] · Anne-Sophie Litmeyer[2] · Jens Huober[7] · Dominik Gerber[8] · Christine Solbach[9] · Vesna Bjelic-Radisic[10,11] · Kai Saeger[12] · Andreas Schneeweiss[13] · Kerstin Rhiem[14] · Moritz Gleitsmann[2] · Susanne von Gerlach[2] · Bärbel Felder[15] · Moritz Jesinghaus[2] · Carsten Denkert[2] · Sibylle Loibl[15] · Wolfgang D. Schmitt[16]

✉ Paul Jank
paul.jank@uni-marburg.de

✉ Carsten Denkert
carsten.denkert@uni-marburg.de

1 Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt Universität Zu Berlin, Department of Gynecology With Breast Center, Berlin, Germany

2 Institute of Pathology, Philipps-University Marburg, Marburg, Germany

3 Helios Klinikum Berlin-Buch, Berlin, Germany

4 Breast Unit, University Hospital Mannheim, Mannheim, Germany

5 Department of Gynecology and Obstetrics, Medical Faculty and University Hospital Carl Gustav CarusTechnische Universität Dresden, Dresden, Germany

6 Rotkreuzklinikum Munich, Munich, Germany

7 HOCH Health Ostschweiz, Kantonsspital St. Gallen, Universitäres Lehr-Und ForschungsspitalBrustzentrum, St. Gallen, Switzerland

8 PreciPoint GmbH, Munich, Germany

9 Breast Unit, Department of Gynecology and Obstetrics, University Hospital Frankfurt, Frankfurt, Germany

10 Helios University Clinic Wuppertal, Wuppertal, Germany

11 University Witten/Herdecke, Witten, Germany

12 VMscope GmbH, Berlin, Germany

13 National Center for Tumor Diseases, University Hospital and German Cancer Research Center, Heidelberg, Germany

14 Zentrum Familiärer Brust-Und EierstockkrebsUniversitätsklinikum Köln, Heidelberg, Germany

15 GBG c/o GBG Forschungs GmbH, Neu-Isenburg, Germany

16 Charité–Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt Universität Zu Berlin, Institute of Pathology, Berlin, Germany