AI-Based Ki67 Quantification for Triage of Genomic Recurrence Score in Breast Cancer

Amanda Dy[1], Abubakr Shafique[1], Xiaoli Qin[1], Dimitrios Androutsos[1], Susan J. Done[2,3,4], and April Khademi[1,5,6,7,8]

[1] Department of Electrical, Computer and Biomedical Engineering, Toronto Metropolitan University, Toronto, ON, Canada
amanda.dy@torontomu.ca
[2] Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada
[3] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada
[4] Laboratory Medicine Program, University Health Network, Toronto, ON, Canada
[5] Vector Institute for Artificial Intelligence, Toronto, ON, Canada
[6] Department of Medical Imaging, University of Toronto, Toronto, ON, Canada
[7] Institute of Medical Science, University of Toronto, Toronto, ON, Canada
[8] Institute for Biomedical Engineering, Science and Technology (iBEST), Toronto, ON, Canada

Abstract. The Oncotype DX Recurrence Score (RS) guides adjuvant chemotherapy decisions in ER+/HER2− early-stage breast cancer, but genomic testing is not universally available. Ki67 reflects tumor proliferation and is biologically associated with RS, yet routine assessment is limited by interobserver variability. We investigated whether independently trained and analytically validated AI-based Ki67 quantification tools provide incremental information beyond routinely available clinicopathological variables for estimating the genomic recurrence score. We examined 167 patients who underwent clinical Oncotype DX testing. Archival tumor tissue was assembled into tissue microarrays, stained for Ki67, and digitized. Three independently developed AI models were analytically validated against reference assessments before RS modeling. Associations with continuous RS were evaluated using linear regression, and logistic regression with cross-validation was used to identify patients unlikely to have RS>25. AI-based Ki67 indices were associated with RS and independently explained one-third of RS variance. When combined with routine clinicopathological variables (age, histological grade, ER, and PR), over half of RS variance was explained, and a 0.90 AUROC was achieved for identifying RS>25. At a threshold of ≥95% sensitivity, the combined variables ruled out 44% of patients, with an NPV of 0.97, compared with 10% and 0.88, respectively, using routine clinicopathological variables alone. Rather than replacing genomic testing, the combined use of AI-based Ki67 quantification and routine clinicopathological variables may serve as an accessible, scalable triage tool to identify patients unlikely to have a high genomic recurrence score. Prospective external validation in whole-slide or biopsy cohorts is required before clinical implementation.

Keywords: Ki67 Index · Oncotype DX Recurrence Score · Breast Cancer.

1 Introduction

Breast cancer is the most common malignancy among women worldwide and remains a leading cause of cancer-related mortality [12]. Within the spectrum of invasive ductal carcinomas, the subset of hormone receptor-positive/HER2-negative (HR+/HER2−) breast cancer represents the largest molecular subtype that is generally associated with a favorable prognosis [1]. Nevertheless, clinical outcomes within this group vary, with a subset of patients experiencing disease recurrence despite standard therapy [15]. Accurate identification of patients at increased risk of recurrence is therefore central to optimizing treatment selection, balancing the benefits of adjuvant chemotherapy against its associated toxicity, cost, and impact on quality of life [18].

To refine recurrence risk estimation in patients with estrogen receptor-positive (ER+)/HER2− breast cancer, multigene expression assays such as the Oncotype DX Breast Recurrence Score (RS), a validated 21-gene expression assay [14], have become integral components of contemporary clinical decision-making by identifying patients most likely to benefit from adjuvant chemotherapy while avoiding overtreatment in those with low-risk disease. However, Oncotype DX testing is centralized and expensive, and its adoption remains limited across health systems. Because cellular proliferation is a major determinant of recurrence risk and constitutes one of the principal biological processes captured by the Oncotype DX assay, the immunohistochemical (IHC) marker Ki67 has consistently demonstrated prognostic value and a significant association with the Oncotype DX RS. Multiple studies have shown that Ki67 is among the strongest pathological correlates of recurrence score, although it is not the sole determinant [17]. However, concerns have also been raised regarding the clinical implementation of Ki67, particularly its interobserver and interlaboratory variability, lack of universally accepted scoring methodologies, and uncertainty surrounding optimal cutoffs, all of which limit its use as a standalone biomarker [6,13]. More recently, studies have highlighted the limitations of relying on Ki67 alone for estimating the Oncotype DX RS [4]. To our knowledge, however, the incremental value of AI-based Ki67 quantification beyond routinely available clinicopathological features for estimating the Oncotype DX RS in an ER+/HER2− patient population has not been comprehensively explored as a triage tool.

Here, we developed and evaluated a set of computational workflows that integrate Ki67 with established clinicopathological variables to estimate the Oncotype DX RS and identify patients unlikely to have a high RS in ER+/HER2− breast cancer. We assessed the relative contribution of individual features and compared model performance across multiple computational approaches. We established a straightforward multivariable and interpretable approach for triaging genomic testing. We highlight the potential of AI-based Ki67 quantification and

routine diagnostic data to serve as an inexpensive, scalable triage tool to identify patients unlikely to have high genomic recurrence scores.

2 Materials and Methods

2.1 Study Cohort

The cohort comprised 167 patients with ER+/HER2− early-stage breast cancer who underwent Oncotype DX testing as part of routine care at Princess Margaret Cancer Centre, Toronto, Canada, between 2008 and 2017 [16]. For each patient, the clinical record supplied the Oncotype DX RS, age, histological grade, and ER and PR expression interpreted according to the American Society of Clinical Oncology and the College of American Pathologists guidelines [11]. Ki67 was not available from the original clinical reports and was measured experimentally. Archival formalin-fixed paraffin-embedded tumor blocks corresponding to the Oncotype-tested cancers were retrieved from Princess Margaret Cancer Centre, and for each patient two representative tissue-microarray tumor cores were constructed, stained for Ki67 and digitized at 20× magnification (0.5 µm/pixel) using an Aperio AT2 scanner. Cores with insufficient tissue were excluded, leaving 268 tumor cores available for analysis.

2.2 Deep Learning Ki67 Quantification and Analytical Validation

Three published AI models for Ki67 quantification were evaluated: piNET [9], UV-Net [7], and DeepLIIF [10]. Each model was previously trained on datasets independent of this study. UV-Net and piNET are multiclass cell-detection models for Ki67 tumor cell detection in breast cancer. They use convolutional encoder–decoders to detect invasive tumor nuclei and classify each nucleus as Ki67 positive or negative. Both models have been evaluated across external breast cancer datasets [8], and piNET was validated as a fully automated Ki67 scoring pipeline [3], whereas UV-Net was used in AI-assisted Ki67 scoring studies [7]. DeepLIIF was included as an IHC quantification comparator; it was not developed specifically for Ki67 or breast cancer. DeepLIIF is a conditional generative adversarial network that segments all nuclei and classifies IHC nuclear staining status and was externally evaluated in breast cancer Ki67 scoring.

For each patient, the Ki67 index was computed across cores and reported as the percentage of Ki67-positive nuclei among the nuclei detected by each model. We analytically validated each model to confirm that its Ki67 index reflected reference assessment. From the cohort, 20 randomly selected cores, one per patient, were reviewed by a board-certified pathologist for invasive tumor-cell count and Ki67 index. Model estimates were compared with this reference using linear regression, separately for tumor-cell count and Ki67 index. Agreement was summarized by the coefficient of determination (R^2) and Lin's concordance correlation coefficient (CCC). Analytical validity was prespecified as $R^2 \geq 0.70$ and CCC ≥ 0.70; models not meeting these criteria were excluded from analysis.

2.3 Statistical Analysis

To examine relationships between variables and recurrence score, univariable associations with recurrence score were examined for AI-based Ki67 index and clinicopathological variables. Continuous associations with recurrence score were assessed using Spearman rank correlation. Differences between risk groups, defined as RS ≤ 25 (low risk) and RS > 25 (high risk) following the Oncotype DX TAILORx trial [19], were evaluated using the Mann-Whitney U test.

Linear regression was used to test whether the AI-based Ki67 index added independent information for continuous recurrence score estimation. Three models were compared: a Ki67-only model, a clinical-only model (age, histological grade, ER, PR), and a combined (age, histological grade, ER, PR, and Ki67) model. Model fit was summarized using adjusted R^2, which penalizes additional predictors, and each variable's independent contribution was quantified by partial R^2, with significance assessed by partial F-tests.

Logistic regression was used to classify RS > 25 from Ki67-only, clinical-only, and combined variables. Patient-level probabilities were generated by leave-one-out cross-validation. The area under the receiver operating characteristic curve (AUROC) was calculated from these cross-validated scores, with a 95% confidence interval (CI) by bootstrap resampling (1,000 replicates), and AUROCs were compared by DeLong tests under Holm correction.

A triage analysis evaluated whether each model could identify patients unlikely to have RS > 25. For each model, the threshold was set to retain at least 95% sensitivity because the costly error in this setting is a missed high-risk patient. Patients below this threshold were classified as rule-out eligible, meaning genomic testing could be deferred. Performance was summarized by the ruled-out fraction and the negative predictive value (NPV), each with a 95% CI by out-of-bag bootstrap resampling. Rule-out-eligible and retained patients were compared across recurrence score and continuous clinical features by the Mann–Whitney U test, and across histological grade by the χ^2 test, with Holm correction. All tests were two-sided at $\alpha = 0.05$.

3 Results and Discussion

3.1 Analytical Validity of Deep Learning Ki67 Quantification

UV-Net and piNET met the prespecified analytical validity criteria against pathologist reference, whereas DeepLIIF did not. The three models produced qualitatively different cell-detection patterns in the study cohort (Fig. 1). UV-Net and piNET showed strong concordance with reference tumor-cell counts (CCC = 0.96 and 0.95, respectively) and reference Ki67 index (CCC = 0.84 for both). In contrast, DeepLIIF showed weak correlation with reference tumor-cell counts ($R^2 = 0.30$) and no concordance (CCC = 0.05). DeepLIIF detected a median of $\sim$2,000 nuclei per core indicating substantial over-detection. UV-Net and piNET were carried forward for recurrence score estimation analysis. We note these models ranked patients similarly (Spearman $\rho = 0.92$), but piNET produced higher estimates, with a mean bias of 5.1 points.

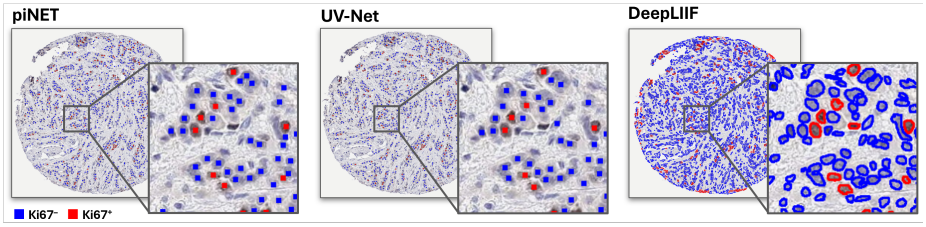

Fig. 1. AI-based Ki67 model outputs on a representative tissue-microarray core.

3.2 Cohort Characteristics and Univariable Associations

Our experimental cohort was enriched for high recurrence score prevalence and showed the expected clinicopathological profile. Of 167 patients, 56 (33.5%) had tumors with RS > 25, compared with 18% in the TAILORx trial [19]. In univariable analyses, RS > 25 tumors were more often grade 3, had lower ER expression, approximately fivefold lower PR expression, and a two- to threefold higher Ki67 index than low recurrence score tumors, regardless of whether Ki67 was measured by UV-Net or piNET (all $p < 0.05$; Table 1). PR expression and AI-based Ki67 showed the strongest associations with continuous recurrence score (absolute Spearman $\rho > 0.5$; Fig. 2). These findings support that AI-based Ki67 captures the proliferation-linked component of recurrence score. However, because Ki67 correlates with clinicopathological features, univariable association alone could not establish independent information. We therefore used multivariable modeling to test whether AI-based Ki67 added recurrence score-related information beyond routine clinical variables.

Table 1. Cohort characteristics by genomic recurrence score groups. Values are median (Q1–Q3), except grade (counts). RS group difference between RS ≤ 25 and RS > 25 by χ^2 (grade) or Mann–Whitney U (otherwise); continuous association by Spearman ρ, grade treated as ordinal.

Variable	RS ≤ 25 n = 111	RS > 25 n = 56	RS group difference	Continuous association
Age, years	56 (49–62)	57 (52–61)	0.81	−0.05
ER, %	99 (95–100)	98 (80–99)	<0.05	−0.28
PR, %	80 (50–90)	15 (0–42)	<0.001	−0.52
Grade, 1/2/3, n	2 / 71 / 38	1 / 24 / 31	<0.05	0.28
Ki67, UV-Net, %	8 (4–12)	23 (12–31)	<0.001	0.52
Ki67, piNET, %	13 (6–16)	29 (18–37)	<0.001	0.54

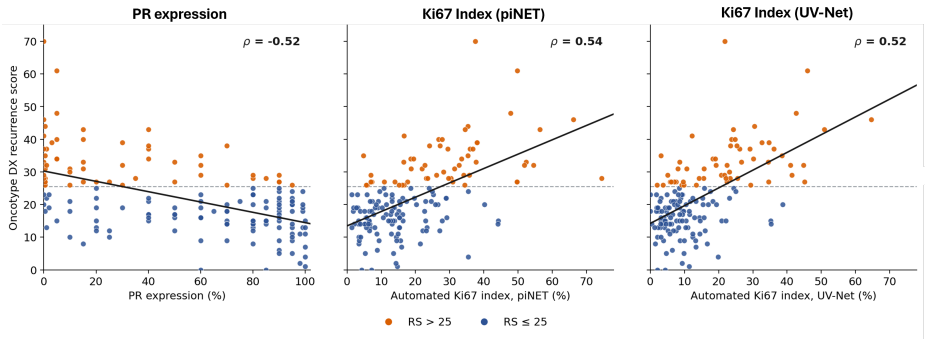

Fig. 2. Strongest univariable correlates of the Oncotype DX recurrence score. PR expression and AI-based Ki67 indices had the largest absolute Spearman correlations (ρ) with recurrence score among candidate variables. Points represent 167 patients; solid lines show linear trends; dashed line marks a recurrence score of 25.

3.3 Multivariable Continuous Recurrence Score Modeling

The AI-based Ki67 index improved continuous recurrence score estimation beyond routine clinicopathological variables and emerged as a main independent contributor (Fig. 3). Ki67 alone explained approximately one-third of recurrence score variance, with adjusted R^2 values of 0.35 for UV-Net and 0.32 for piNET, indicating that Ki67 captured substantial recurrence score information on its own. The clinical-only model explained approximately half of recurrence score variance (adjusted $R^2 = 0.47$), with PR expression, ER expression, and histological grade contributing independently. Adding AI-based Ki67 index to routine clinical variables increased adjusted R^2 to 0.57 for UV-Net and 0.56 for piNET. In both combined models, Ki67 was the second-largest independent contributor after PR. Its inclusion also changed the contribution of histological grade. Grade contributed independently in the clinical-only model but no longer contributed after Ki67 was added, suggesting that AI-based Ki67 captured the proliferation component that histological grade reflects more coarsely through mitotic activity. Together, these findings identify AI-based Ki67 as an interpretable and independent source of recurrence score information beyond routine clinicopathological variables.

3.4 Recurrence Score Discrimination and Triage Analysis

For the clinically relevant task of identifying high genomic recurrence score, the independent information captured by AI-based Ki67 indices translated into improved discrimination (Fig. 4). The combined models achieved the highest AUROCs, with values of 0.90 for UV-Net and 0.88 for piNET when combined with routine clinicopathological variables. By comparison, the clinical-only model and Ki67 alone each achieved an AUROC of 0.81. Both combined models improved discrimination compared with the clinical-only model (UV-Net: $p < 0.01$; piNET:

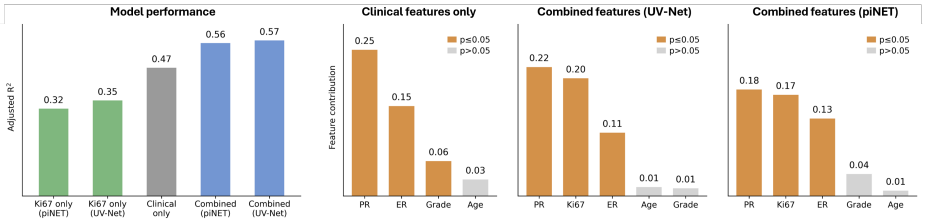

Fig. 3. Ki67 contributes independently to continuous recurrence score estimation. Feature contributions are shown as partial R^2, indicating the independent variance explained by each predictor.

$p < 0.05$), whereas Ki67-only and clinical-only models did not differ ($p = 1.00$). These results show that AI-based Ki67 index matched the routine clinicopathological variables as a standalone discriminator but added independent discriminatory information when combined with them.

At a high-sensitivity operating point, adding the AI-based Ki67 index expanded the population ruled out for high genomic recurrence score triage (Fig. 4). When thresholds were set to maintain at least 95% sensitivity for RS > 25, the combined models ruled out 44% of patients (95% CI, 18–67%) with an NPV of 0.97 (95% CI, 0.85–1.00). By comparison, the clinical-only model ruled out 10% of patients (95% CI, 7–45%) with an NPV of 0.88 (95% CI, 0.75–1.00). Among patients that would require genomic testing, the proportion with RS > 25 was also higher for the combined models than for the clinical-only model, corresponding to a positive predictive value (PPV) of 0.60 (95% CI, 0.40–0.78) versus 0.40 (95% CI, 0.28–0.53). The combined model therefore sent fewer patients for genomic testing, a greater share of whom truly had a high recurrence score, and it missed no additional high recurrence score patients.

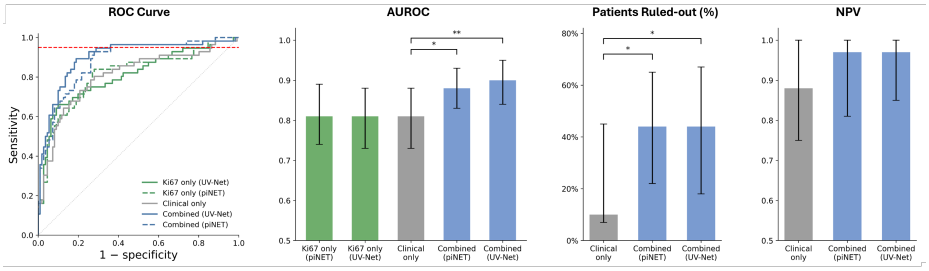

Fig. 4. Ki67 improves high recurrence score discrimination and rule-out triage performance at the ≥95% sensitivity threshold. Receiver operating characteristic (ROC) curves show discrimination of RS > 25. The red dashed line indicates the 95% sensitivity operating point for triage analysis. Bar plots show mean and 95% CI. *$p < 0.05$; **$p < 0.01$.

The reported negative predictive value is likely conservative, because this cohort was enriched for high genomic recurrence score at 33.5% against roughly 18% in screening populations [19]. Sensitivity was calculated among high recurrence score patients, so the prespecified ≥95% sensitivity target was not determined by this enrichment. The negative predictive value, in contrast, depends on prevalence and would be expected to rise in a lower-prevalence intended-use population if sensitivity and specificity were preserved. The ruled-out fraction may also increase in such a population, but this estimate should be confirmed in an external intended-use cohort.

The population cleared by the threshold was biologically plausible, showing higher PR expression and lower AI-based Ki67 than the retained patients (Fig. 5). PR expression in the combined models was 90% versus 20% and the clinical-only model 95% versus 50%. The Ki67 index was 6% versus 17% in the combined UV-Net model and 9% versus 24% in the combined piNET model. The pattern is biologically coherent, because a high Ki67 index reflects both high proliferation and recurrence risk [17], and low PR expression has been associated with high recurrence score [20] and adverse outcomes in ER-positive disease [2]. This pattern may hold beyond this cohort, although the internally derived threshold would require recalibration in a new population.

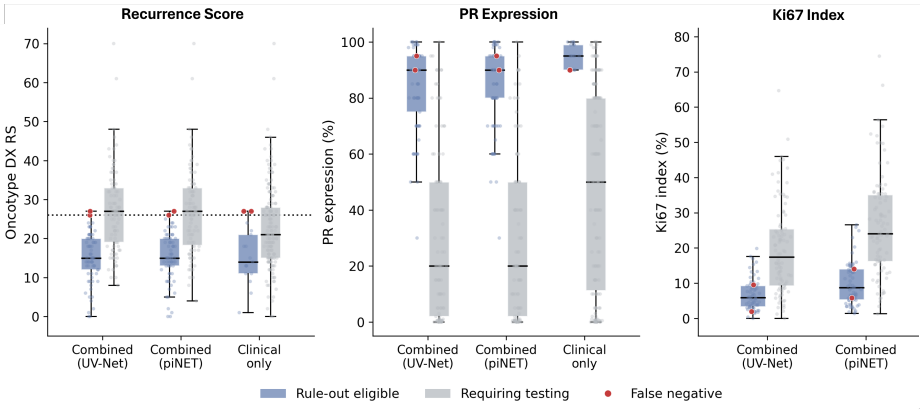

Fig. 5. Distributions of rule-out-eligible versus requiring-testing patients at the ≥95% sensitivity threshold. Boxplots show median and interquartile range; points show patients. The dotted line marks RS = 26. All group differences were significant ($p < 0.001$).

Several limitations constrain these results. The study was retrospective, drawn from a single institution, and built on tissue-microarray cores, which standardize staining and analysis but may undersample the intratumoral heterogeneity that larger tumor sections would capture. Long-term recurrence outcomes were not available, so the recurrence score served as the prediction target in place of

disease recurrence itself, and the models therefore estimate genomic recurrence score rather than clinical outcome. The rule-out threshold was derived and evaluated within one cohort and rested on a small number of high recurrence score patients, which makes its operating characteristics exploratory and the confidence intervals wide. The analysis also depended on one to two representative tissue microarray (TMA) cores per patient rather than the whole-slide section or core-needle biopsy that a diagnostic workflow would use. Prospective validation in whole-slide or biopsy cohorts that assess real-world utility, reproducibility, and impact is required before routine adoption [5].

4 Conclusion

Analytically validated AI-based Ki67 quantification provided incremental information beyond routine clinicopathological variables for estimating Oncotype DX RS in ER+/HER2− early-stage breast cancer. When combined with age, histological grade, ER, and PR, AI-based Ki67 improved continuous recurrence score estimation, discrimination of RS > 25, and high-sensitivity rule-out triage compared with clinicopathological variables alone. These findings support AI-based Ki67 quantification as a potential pre-genomic testing triage tool to identify patients unlikely to have high genomic recurrence scores. Prospective and external validation in intended-use whole-slide or biopsy cohorts is required before clinical implementation.

Acknowledgments. A.D. was supported by the Natural Sciences and Engineering Research Council of Canada Vanier Canada Graduate Scholarship (198875). This work was funded by the Canadian Cancer Society (707155) and Mitacs Accelerate, Canada (IT33068). The funders had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Burstein, H.J., Curigliano, G., Loibl, S., Dubsky, P., Gnant, M., Poortmans, P., Colleoni, M., Denkert, C., Piccart-Gebhart, M., Regan, M., Senn, H.J., Winer, E.P., Thurlimann, B.: Estimating the benefits of therapy for early-stage breast cancer: the St. Gallen International Consensus Guidelines for the primary therapy of early breast cancer 2019. Ann. Oncol. **30**(10), 1541–1557 (2019)
2. Clark, B.Z., Dabbs, D.J., Cooper, K.L., Bhargava, R.: Impact of progesterone receptor semiquantitative immunohistochemical result on Oncotype DX recurrence score: a quality assurance study of 1074 cases. Appl. Immunohistochem. Mol. Morphol. **21**(4), 287–291 (2013)
3. Dawe, M., Shi, W., Liu, T.Y., Lajkosz, K., Shibahara, Y., Gopal, N.E.K., Geread, R., Mirjahanmardi, S., Wei, C.X., Butt, S., Abdalla, M., Manolescu, S., Liang, S.B., Chadwick, D., Roehrl, M.H.A., McKee, T.D., Adeoye, A., McCready, D.,

Khademi, A., Liu, F.F., Fyles, A., Done, S.J.: Reliability and variability of Ki-67 digital image analysis methods for clinical diagnostics in breast cancer. Lab. Invest. **104**(5), 100341 (2024)

4. Dragoumis, D., Kapetsis, G., Louis, K., Maniatis, D., Mpalampou, E., Bouloukos, K., Xenakis, X., Papaioannou, N., Parpoudi, S., Pesmatzoglou, G., Sachoulidou, A., Sfakianakis, E., Triantafyllidou, S., Tsantilas, V., Tsiftsoglou, A., Filippidou, S., Fyssas, I., Stathoulopoulou, M., Matiatou, M., Karathanasis, P., Alexandrou, D., Amanatidou, A., Desiris, K., Efraimidou, E., Zavos, A., Michailidou, E., Roussogiannis, S., Venizelos, V.: Evaluating the association of Ki-67 with Oncotype DX Recurrence Score in early-stage ER-positive/HER2-negative breast cancer. Cancers **18**(11), 1731 (2026)
5. Dy, A., Buetow, S.M., Bredemeyer, A.J., Lamba Saini, M., Lucas, F., Bennett, S., Blenman, K.R.M., Wharton Jr., K., Singhal, S., de Baca, M.E., Schap, K., Hanna, M.G., Kearney, S.J., Zerbe, N., Salgado, R., Veetil, J., Seheult, J.N., McClintock, D.S., Khademi, A., Lennerz, J.K.: Clarifying validation terminologies in healthcare. npj Digit. Med. (2026). https://doi.org/10.1038/s41746-026-02471-2
6. Dy, A., Lennerz, J.K.: Ki-67 as a treatment decision modifier in breast cancer. Oncologist **31**(7), oyag183 (2026)
7. Dy, A., Nguyen, N.N.J., Meyer, J., Dawe, M., Shi, W., Androutsos, D., Fyles, A., Liu, F.F., Done, S., Khademi, A.: AI improves accuracy, agreement and efficiency of pathologists for Ki67 assessments in breast cancer. Sci. Rep. **14**, 1283 (2024)
8. Dy, A., Nguyen, N.N.J., Mirjahanmardi, S.H., Androutsos, D., Dawe, M., Fyles, A., Shi, W., Liu, F.F., Done, S., Khademi, A.: Domain adaptation using silver standard labels for Ki-67 scoring in digital pathology: a step closer to widescale deployment. In: Medical Imaging with Deep Learning. PMLR, vol. 227, pp. 653–665 (2024)
9. Geread, R.S., Sivanandarajah, A., Brouwer, E.R., Wood, G.A., Androutsos, D., Faragalla, H., Khademi, A.: piNET: an automated proliferation index calculator framework for Ki67 breast cancer images. Cancers **13**(1), 11 (2021)
10. Ghahremani, P., Li, Y., Kaufman, A., Vanguri, R., Greenwald, N.F., Angelo, M., Hollmann, T.J., Nadeem, S.: Deep learning-inferred multiplex immunofluorescence for immunohistochemical image quantification. Nat. Mach. Intell. **4**(4), 401–412 (2022)
11. Hammond, M.E.H., Hayes, D.F., Dowsett, M., Allred, D.C., Hagerty, K.L., Badve, S., Fitzgibbons, P.L., Francis, G., Goldstein, N.S., Hayes, M., Hicks, D.G., Lester, S., Love, R., Mangu, P.B., McShane, L.M., Miller, K., Osborne, C.K., Paik, S., Perlmutter, J., Rhodes, A., Sasano, H., Schwartz, J.N., Sweep, F.C.G., Taube, S., Torlakovic, E.E., Valenstein, P., Viale, G., Visscher, D., Wheeler, T., Williams, R.B., Wittliff, J.L., Wolff, A.C.: American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J. Clin. Oncol. **28**(16), 2784–2795 (2010)
12. Hankinson, S.E., Polyak, K., Garber, J.E., Anderson, B.O.: Breast cancer: multiple, often complex, risk factors. In: Wild, C.P., Weiderpass, E., Stewart, B.W. (eds.) World Cancer Report: Cancer Research for Cancer Prevention, Section 5.9. International Agency for Research on Cancer, Lyon (2020)
13. Nielsen, T.O., Leung, S.C.Y., Rimm, D.L., Dodson, A., Acs, B., Badve, S., Denkert, C., Ellis, M.J., Fineberg, S., Flowers, M., Kreipe, H.H., Laenkholm, A.V., Pan, H., Penault-Llorca, F.M., Polley, M.Y., Salgado, R., Smith, I.E., Sugie, T., Bartlett, J., McShane, L.M., Dowsett, M., Hayes, D.F.: Assessment of Ki67 in breast cancer:

updated recommendations from the International Ki67 in Breast Cancer Working Group. J. Natl. Cancer Inst. **113**(7), 808–819 (2021)
14. Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., Baehner, F.L., Walker, M.G., Watson, D., Park, T., Hiller, W., Fisher, E.R., Wickerham, D.L., Bryant, J., Wolmark, N.: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N. Engl. J. Med. **351**(27), 2817–2826 (2004)
15. Pan, H., Gray, R., Braybrooke, J., Davies, C., Taylor, C., McGale, P., Peto, R., Pritchard, K.I., Bergh, J., Dowsett, M., Hayes, D.F.: 20-year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years. N. Engl. J. Med. **377**(19), 1836–1846 (2017)
16. Rotbauer, M., Dawe, M., Nixon, K.C.J., Tsao, J., Durbic, T., Sharma, M., Liang, S.B., Zhang, J., Gopal, N.E.K., Mrkonjic, M., Alshamlan, N., Karavelic, A., Murray, C., Cescon, D.W., Bedard, P.L., Done, S.J.: The role of circulating tumor cell-associated genes in the progression of estrogen receptor-positive breast cancer. npj Breast Cancer **12**, 11 (2026). https://doi.org/10.1038/s41523-025-00874-0
17. Sahebjam, S., Aloyz, R., Pilavdzic, D., Brisson, M.L., Ferrario, C., Bouganim, N., Cohen, V., Miller Jr., W.H., Panasci, L.C.: Ki 67 is a major, but not the sole determinant of Oncotype DX recurrence score. Br. J. Cancer **105**, 1342–1345 (2011)
18. Salvo, E.M., Ramirez, A.O., Cueto, J., Law, E.H., Situ, A., Cameron, C., Samjoo, I.A.: Risk of recurrence among patients with HR-positive, HER2-negative, early breast cancer receiving adjuvant endocrine therapy: a systematic review and meta-analysis. Breast **57**, 5–17 (2021)
19. Sparano, J.A., Gray, R.J., Makower, D.F., Pritchard, K.I., Albain, K.S., Hayes, D.F., Geyer Jr., C.E., Dees, E.C., Goetz, M.P., Olson Jr., J.A., Lively, T., Badve, S.S., Saphner, T.J., Wagner, L.I., Whelan, T.J., Ellis, M.J., Paik, S., Wood, W.C., Ravdin, P.M., Keane, M.M., Gomez Moreno, H.L., Reddy, P.S., Goggins, T.F., Mayer, I.A., Brufsky, A.M., Toppmeyer, D.L., Kaklamani, V.G., Atkins, J.N., Berenberg, J.L., Sledge, G.W.: Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. N. Engl. J. Med. **379**(2), 111–121 (2018)
20. Sughayer, M.A., Maraqa, B., Qura'an, B., Alsughayer, A., Abdel-Razeq, H.: A simplified novel algorithm to predict the 21-gene recurrence score. World J. Oncol. **16**(6), 555–564 (2025)