ScientificScholar®
Knowledge is power
Publisher of Scientific Journals

www.ijmio.com

International Journal of Molecular and Immuno Oncology

Article in Press

Original Article

Comparison of three counting methods for Ki-67 labeling index in breast carcinoma: A study on core biopsies and resection specimens

Satyajit Singh Gill[1], Nutan Jumle[1], Gourav Agrawal[1], Shrikanth Muralidharan[2]

[1]Department of Pathology and Laboratory Services, [2]Department of Medical Administration, Jehangir Hospital, Pune, Maharashtra, India.

***Corresponding author:**
Shrikanth Muralidharan, Department of Medical Administration, Jehangir Hospital, Pune, Maharashtra, India.

shrikanth.muralidharan@jehangirhospital.com

Received: 20 February 2026
Accepted: 11 May 2026
EPub Ahead of Print: 07 July 2026
Published:

DOI
10.25259/IJMIO_7_2026

Quick Response Code:

ABSTRACT

Objectives: Ki-67 labeling index (LI) is a crucial prognostic and predictive biomarker in breast carcinoma, but its clinical utility is limited by a lack of standardization in scoring methods. This study evaluates the hotspot, average, and tumor edge methods for Ki-67 LI assessment in core needle biopsies and resection specimens.

Material and Methods: A combined retrospective and prospective observational study was conducted on 200 cases of primary invasive breast carcinoma. Ki-67 LI was determined using immunohistochemistry and counted through the three methods. Correlations with histological grade, estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status were analyzed.

Results: Mean Ki-67 LI was highest with the hotspot method (51.9% in biopsies, 50.5% in resections), followed by tumor edge (46.8%, 46.9%) and average (44.5%, 40.3%) methods ($P < 0.001$ for inter-method differences). All methods correlated significantly with higher histological grades and negative ER/PR/HER2 status ($P < 0.001$). Intraclass correlation coefficient values exceeded 0.9, indicating strong concordance, with hotspot and tumor edge methods showing the highest agreement. Categorization into low (<20%) and high (≥20%) Ki-67 groups revealed minimal differences in biopsies but a higher prevalence of low-Ki-67 tumors with the average method in resections ($P < 0.05$).

Conclusion: The hotspot method yields consistently higher Ki-67 LI and better concordance across specimen types.

Keywords: Average method, Breast carcinoma, Hotspot method, Immunohistochemistry, Ki-67 labeling index, Tumor edge method

INTRODUCTION

Breast cancer remains the most prevalent malignancy among women globally, with an estimated 2.1 million new diagnoses and over 600,000 deaths.[1] Effective management hinges on accurate prognostic and predictive factors, including lymph node status, tumor size, histological grade, and molecular biomarkers such as hormone receptors (estrogen receptor [ER] and progesterone receptor [PR]), human epidermal growth factor receptor 2 (HER2), and Ki-67 proliferation index. Ki-67, a nuclear protein expressed in proliferating cells, serves as a key marker for tumor proliferation. Its labeling index (LI) informs prognosis, particularly in conjunction with nodal status, tumor size, and receptor profiles.[2] Despite two decades of use, Ki-67 integration into

clinical practice is hampered by scoring variability due to tumor heterogeneity and methodological differences. Proposed counting methods include hotspot (focusing on areas of highest proliferation), average (random fields), and tumor edge (invasive fronts).[3] Tissue type – core needle biopsy versus resection specimen – further influences method choice, as biopsies often provide limited material.[4] This study aimed to compare these three methods in core biopsies and resection specimens, correlate Ki-67 LI with histological grade and receptor status, assess inter-method concordance, and recommend appropriate approaches for each specimen type.

MATERIAL AND METHODS

Study design and population

This was a combined retrospective and prospective observational study approved by the Institutional Review Board of Jehangir Hospital, Pune. Inclusion criteria encompassed all primary invasive breast carcinoma cases diagnosed through core biopsies or surgical resections in the year 2023–2024. Exclusion criteria included *in situ* carcinomas and post-neoadjuvant chemotherapy specimens. A sample size of 200 cases was involved.

Data collection and processing

Clinical, gross, and microscopic features were recorded per a standardized pro forma. Surgical specimens were fixed in 10% neutral-buffered formalin and paraffin embedded. ER, PR, and HER2 status were evaluated by immunohistochemistry (IHC) at diagnosis, following 2010 and 2013 American Society of Clinical Oncology/College of American Pathologists guidelines with 2018 updates.[5] ER/PR positivity was defined as ≥1% nuclear staining, and HER2 scoring followed established criteria. For Ki-67 IHC, blocks were selected for maximum tumor area (core biopsies) or tumor with adjacent normal tissue (resections). Four-micrometer sections underwent heat-induced antigen retrieval and staining with rabbit monoclonal antibody SP-6 (1:100; Thermo Scientific). Tonsil sections served as external controls, with lymphocytes and benign ducts as internal controls.

Ki-67 interpretation

Fields were selected by a single observer to minimize variability.[6] Counting occurred at ×400 magnification (field diameter 0.4 mm) using a Leica DMLB microscope. Nuclear staining was considered positive, regardless of intensity; cytoplasmic/membranous staining was excluded. At least 500 tumor cells were counted per case, adjusted for low-density samples.

Ki-67 LI was calculated as follows:

- Average method: Three random fields (varied density preferred); positive cells averaged as a percentage
- Hotspot method: Hottest area under ×100; positive cells counted at ×400
- Tumor edge method: Invasive edge or stromal-tumor interface; positive cells counted at ×400.[7]

Categorical data were presented as frequencies/percentages, and continuous data were presented as mean ± standard deviation. $P < 0.05$ indicated statistical significance. Analyses were performed using the Statistical Package for the Social Sciences version 21.0 (IBM Corporation, USA).

RESULTS

Demographic and clinicopathological characteristics

The cohort comprised 200 cases (mean age 54.9 ± 12.7 years, range 28–90). The 41–50 age group was the most common (25.0%). Right-sided tumors predominated (51.5%). Specimens included 98 core biopsies (49.0%) and 102 resections (51.0%). Histological grades were grade 1 (8.5%), grade 2 (51.5%), and grade 3 (40.0%). In core biopsies [Table 1], mean Ki-67 LI increased with grade across methods ($P < 0.001$ inter-grade).

Hotspot yielded the highest values (overall 51.9% ± 21.1%), followed by tumor edge (46.8% ± 21.1%) and average (44.5% ± 18.2%) ($P < 0.001$ inter-method). Differences were non-significant in grade 1 but significant in grades 2–3 ($P < 0.01$). In resections [Table 2], similar trends emerged: Hotspot (50.5% ± 23.4%), tumor edge (46.9% ± 22.6%), and average (40.3% ± 20.6%) ($P < 0.001$).

Table 1: Mean Ki-67 LI by method and grade in core biopsies.

Grade	Hotspot (mean±SD)	Tumor edge (mean±SD)	Average (mean±SD)
1	32.5±24.3	29.1±23.4	25.5±13.6
2	52.2±20.0	48.1±20.6	44.3±17.9
3	58.9±17.1	55.0±14.6	51.8±15.2
Overall	51.9±21.1	46.8±21.1	44.5±18.2

SD: Standard deviation, LI: Labeling index

Table 2: Mean Ki-67 LI by method and grade in resections.

Grade	Hotspot (mean±SD)	Tumor edge (mean±SD)	Average (mean±SD)
1	36.2±26.9	35.7±27.5	24.7±18.4
2	41.7±22.4	38.6±21.3	33.1±19.1
3	62.3±19.1	58.1±18.9	50.1±18.4
Overall	50.5±23.4	46.9±22.6	40.3±20.6

SD: Standard deviation, LI: Labeling index

Inter-grade differences were significant ($P < 0.001$), with method variances notable in grades 2–3. $P < 0.001$ for inter-method and inter-grade comparisons (except grade 1 inter-method, $P > 0.05$). In biopsies [Table 3], Ki-67 LI was higher in ER/PR/HER2-negative cases ($P < 0.001$). Hotspot showed the greatest differences (e.g., 39.1% change ER-positive vs. negative).

In resections [Table 4], similar patterns held, with larger percentage changes (e.g., 45.1% for ER).

Using a 20% cutoff, high Ki-67 (≥20%) predominated (90–94%) based on previous published literature.[8] Differences were non-significant in biopsies but significant in resections for average versus others ($P < 0.05$). ICC values were >0.9, with the hotspot-tumor edge highest (0.948–0.964).

DISCUSSION

The evaluation of Ki-67 LI in breast carcinoma remains a cornerstone for assessing tumor proliferation, yet its integration into routine clinical decision-making is fraught with challenges stemming from methodological inconsistencies and interpretive variability.[6,7,9] In this study, we systematically compared three manual counting methods – hotspot, average, and tumor edge – across core needle biopsies and resection specimens in 200 cases of invasive breast carcinoma. Our findings reveal that the hotspot method consistently yields the highest mean Ki-67 LI (51.9% in biopsies and 50.5% in resections), followed by the tumor edge (46.8% and 46.9%) and average methods (44.5% and 40.3%), with statistically significant inter-method differences. These results align with prior investigations demonstrating that hotspot scoring captures areas of highest proliferation, often resulting in elevated indices compared to averaging approaches that dilute heterogeneous expression.[10] Critically, the observed differences underscore the inherent subjectivity in manual Ki-67 assessment. The hotspot method, while intuitive for identifying aggressive clones, relies on observer judgment to select the "hottest" area, introducing bias, particularly in heterogeneous tumors like breast carcinoma.[11] In contrast, the average method, by sampling random fields, aims for representativeness but may underestimate proliferation in tumors with focal high-activity regions, as evidenced by its significantly lower values in higher-grade lesions. The tumor edge method, focusing on invasive fronts, offers a biologically plausible alternative, potentially capturing cells with enhanced migratory and proliferative potential, which could correlate better with metastatic risk.[12] Yet, its applicability is limited in core biopsies lacking clear edges, where the stromal-tumor interface was selected for assessment, potentially confounding results. Our data show that inter-method variances are more pronounced in resection specimens, where larger tissue volumes amplify heterogeneity, suggesting that method choice must be specimen-specific to avoid discordant prognostic estimates. Correlations with histological grade and receptor status further illuminate the clinical relevance of these methods. Across all approaches, Ki-67 LI increased significantly with grade, reinforcing its role as a surrogate for tumor aggressiveness.[4] Notably, higher LI in grade 3 tumors highlights proliferation as a driver of poor differentiation. Similarly, inverse associations with ER/PR positivity and direct links with HER2 positivity align with molecular subtyping, where luminal A tumors exhibit low proliferation versus triple-negative or HER2-enriched subtypes. However, the percentage changes were method-dependent, with hotspot amplifying differences, which could skew subtype classification and therapeutic decisions. For instance, in ER-positive cases, where Ki-67 informs chemotherapy avoidance, an inflated hotspot score might unnecessarily escalate treatment, contradicting evidence from trials like monarchE, where precise cutoffs are pivotal.[13] This discrepancy emphasizes the need for method-calibrated thresholds, as universal application risks misclassification. Intraclass correlation coefficients (ICC > 0.9) indicate strong concordance among methods, with hotspot and tumor edge showing the highest agreement (0.948–0.964). This suggests reliability within a controlled, single observer setting, but

Table 3: Mean Ki-67 LI by method and receptor in biopsies.

Receptor	Status	Hotspot (mean±SD)	Tumor edge (mean±SD)	Average (mean±SD)
ER	Positive	44.5±22.3	38.1±22.1	37.1±18.1
	Negative	61.9±14.6	59.0±11.9	54.4±13.2
PR	Positive	39.7±21.7	32.7±22.1	32.9±16.5
	Negative	61.2±15.4	56.9±13.5	53.2±14.4
HER2	Positive	58.3±16.3	54.6±18.1	50.1±15.8
	Negative	49.5±22.3	44.6±21.6	42.4±18.8

SD: Standard deviation, LI: Labeling index, ER: Estrogen receptor, PR: Progesterone receptor, HER2: Human epidermal growth factor receptor 2

Table 4: Mean Ki-67 LI by method and receptor in resections.

Receptor	Status	Hotspot (mean±SD)	Tumor edge (mean±SD)	Average (mean±SD)
ER	Positive	42.8±23.8	39.0±22.3	32.7±19.3
	Negative	62.1±17.5	58.6±17.4	51.7±17.1
PR	Positive	39.7±22.1	35.3±19.9	29.5±17.0
	Negative	59.4±20.7	56.3±20.2	49.2±19.1
HER2	Positive	47.4±21.3	44.4±20.4	39.1±19.8
	Negative	51.6±24.1	47.7±23.3	40.7±20.9

SD: Standard deviation, LI: Labeling index, ER: Estrogen receptor, PR: Progesterone receptor, HER2: Human epidermal growth factor receptor 2

such figures are artificially inflated compared to multi-center studies reporting ICCs as low as 0.59 due to inter-observer variability.[7] In resections, tumor edge may better reflect invasive behavior, though the average could suffice for global assessment per some guidelines. However, without harmonization, Ki-67's clinical utility remains suboptimal, as evidenced by its exclusion from some prognostic tools like Oncotype DX due to variability.[14] Digital image analysis (DIA) emerges as a superior alternative, as it identifies more high-Ki-67 cases and reduces variability by automating hotspot detection and cell segmentation. Yet, DIA's adoption is hindered by cost and validation needs, and our manual approach, while cost-effective, perpetuates the "Ki-67 dilemma" of prognostic promise versus practical pitfalls.[15] Specimen type introduces additional critique. Core biopsies yielded higher overall LI than resections (e.g., average 44.5% vs. 40.3%), attributing this to sampling bias toward proliferative centers in limited tissue. This discordance (median difference ~5%) could lead to prognostic mismatches, especially in neoadjuvant settings where biopsy Ki-67 guides therapy but resection reflects post-treatment changes (excluded here). Manual counting's limitations – time consumption (often >30 min/case), fatigue-induced errors, and inconsistent field selection – undermine reproducibility, particularly in routine pathology where workload pressures exacerbate inconsistencies.[16] Moreover, our use of a fixed 500-cell count, while adhering to guidelines, falls short in low-density samples, where undercounting biases low-proliferative tumors. Our exclusion of post-neoadjuvant cases is a strength, avoiding confounding, but it limits generalizability to treated populations. Categorization using a 20% cutoff – a pragmatic but arbitrary threshold – showed minimal inter-method differences in biopsies but higher low-Ki-67 prevalence with average in resections, questioning its utility without outcome data. The international Ki67 in breast cancer working group advocates ≤5% low and ≥30% high for early-stage prognosis, yet our 20% aligns with older practices and may misclassify borderline cases.[15] Critically, without survival follow-up, our correlations remain associative, not predictive; prospective validation is essential to ascertain which method best forecasts recurrence or response.

The single-center, observational design (n = 200) provides robust internal validity but risks selection bias, with overrepresentation of grade 2 tumors (51.5%) potentially skewing toward intermediate proliferation. Single-observer scoring mitigates variability but overestimates real-world reproducibility, where inter-pathologist discordance reaches 20–30.[6] Future directions must prioritize DIA integration, with machine learning models outperforming manual methods in reproducibility and speed.[17] Multicenter trials correlating method-specific LI with outcomes, incorporating AI, are imperative to resolve discrepancies. Until then, pathologists should report the method used, with caveats on interpretive limitations, to inform multidisciplinary decisions.

CONCLUSION

This study underscores the value of Ki-67 LI as a prognostic biomarker in breast carcinoma, demonstrating strong correlations with histological grade and receptor status across all three counting methods. However, given the inherent variability and limitations of manual assessment, the hotspot method emerges as the most reliable and promising approach among the three, offering higher labeling indices, better concordance across specimen types, and enhanced sensitivity to tumor heterogeneity – making it particularly suitable for core biopsies and a viable interim standard for resections until DIA and comprehensive standardization protocols are widely implemented to ensure reproducibility and clinical utility.

Ethical approval: The research/study was approved by the Institutional Review Board at Jehangir Clinical Development Centre Pvt. Ltd., number JCDC/2022/Oct/p05, dated 13th December, 2022.
Declaration of patient consent: The authors certify that they have obtained all appropriate patient consent forms. In the form, the patient has given consent for clinical information to be reported in the journal. The patient understands that the patient's names and initials will not be published and due efforts will be made to conceal their identity, but anonymity cannot be guaranteed.
Financial support and sponsorship: Nil.
Conflicts of interest: There are no conflicts of interest.
Use of artificial intelligence (AI)-assisted technology for manuscript preparation: The authors confirm that they have used artificial intelligence (AI)-assisted technology solely for language refinement and to improve the clarity of writing. No AI assistance was employed in the generation of scientific content, data analysis or interpretation.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Petrelli F, Viale G, Cabiddu M, Barni S. Prognostic value of different cut-off levels of Ki-67 in breast cancer: A systematic review and meta-analysis of 64,196 patients. Breast Cancer Res Treat 2015;153:477-91.
3. Leung SC, Nielsen TO, Zabaglo L, Arun I, Badve SS, Bane AL, *et al*. Analytical validation of a standardized scoring protocol for Ki67: Phase 3 of an international multicenter collaboration. NPJ Breast Cancer 2016;2:16014.
4. Denkert C, Budczies J, Von Minckwitz G, Wienert S, Loibl S, Klauschen F. Strategies for developing Ki67 as a useful biomarker in breast cancer. Breast 2015;24 Suppl 2:S67-72.
5. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, *et al*. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American society

of clinical oncology/college of American pathologists clinical practice guideline update. J Clin Oncol 2013;31:3997-4013.
6. Shui R, Yu B, Bi R, Yang F, Yang W. An interobserver reproducibility analysis of Ki67 visual assessment in breast cancer. PLoS One 2015;10:e0125131.
7. Rimm DL, Leung SC, McShane LM, Bai Y, Bane AL, Bartlett JM, *et al*. An international multicenter study to evaluate reproducibility of automated scoring for assessment of Ki67 in breast cancer. Mod Pathol 2019;32:59-69.
8. Dokcu Ş, Ali-Çaparlar M, Çetindağ Ö, Hakseven M, Eroğlu A. Prognostic value of KI-67 proliferation index in luminal breast cancers. Cir Cir 2023;91:1-8.
9. Muftah AA, Aleskandarany MA, Al-Kaabi MM, Sonbul SN, Diez-Rodriguez M, Nolan CC, *et al*. Ki67 expression in invasive breast cancer: The use of tissue microarrays compared with whole tissue sections. Breast Cancer Res Treat 2017;164:341-8.
10. Jang MH, Kim HJ, Chung YR, Lee Y, Park SY. A comparison of Ki-67 counting methods in luminal Breast Cancer: The average method vs. The hot spot method. PLoS One 2017;12:e0172031.
11. Chung YR, Jang MH, Park SY, Gong G, Jung WH. Interobserver variability of Ki-67 measurement in breast cancer. J Pathol Transl Med 2016;50:129-37.
12. Kadivar M, Aram F. Assessment of Ki67 in breast cancer: A comparison between the eye-10 method, stepwise counting strategy, and international system of Ki67 evaluation. Iranian J Pathol 2020;15:13-8.
13. Harbeck N, Rastogi P, Martin M, Tolaney SM, Shao ZM, Fasching PA, *et al*. Adjuvant abemaciclib combined with endocrine therapy for high-risk early breast cancer: Updated efficacy and Ki-67 analysis from the monarchE study. Ann Oncol 2021;32:1571-81.
14. Furrer D, Sanschagrin F, Jacob S, Diorio C. Low correlation between Ki67 assessed by qRT-PCR in Oncotype DX and IHC in breast cancer tissues. Sci Rep 2022;12:3593.
15. Khan AM, Sirinukunwattana K, Rajpoot N. An advanced deep learning approach for Ki-67 stained hotspot detection and proliferation rate scoring for prognostic evaluation of breast cancer. Sci Rep 2017;7:4103.
16. Linder N, Taylor PR, Collan Y, Kronqvist P, Giacchetti S, Löffler M, *et al*. The Ki67 dilemma: Investigating prognostic cut-offs and reproducibility for automated Ki67 scoring in breast cancer. Breast Cancer Res Treat 2024;206:385-95.
17. Stålhammar G, Fuentes Martinez N, Lippert M, Tobin NP, Mølholm I, Kis L, *et al*. Digital image analysis outperforms manual biomarker assessment in breast cancer. Mod Pathol 2016;29:318-29.

How to cite this article: Gill S, Jumle N, Agrawal G, Muralidharan S. Comparison of three counting methods for Ki-67 labeling index in breast carcinoma: A study on core biopsies and resection specimens. Int J Mol Immuno Oncol. doi: 10.25259/IJMIO_7_2026