INTRODUCTION

Immunohistochemical (IHC) assessment of Ki-67 proliferative index (PI) has been established as a valuable prognostic marker in early breast cancer; however, its role in predicting response to neoadjuvant chemotherapy is still uncertain. Lack of an accepted validated method of assessment of Ki-67 and wide variations in the cut-offs has severely limited its clinical utility. The International Breast Cancer Working Group (IBCWG) recently proposed guidelines for the analysis, reporting and use of Ki-67 to reduce inter-laboratory variability and improve inter-study comparability of Ki-67 results. In a follow-up study, IBCWG found substantial variability in Ki-67 scoring among some of the world’s most experienced laboratories. Factors contributing to inter-laboratory discordance included tumor region selection, counting method, and subjective assessment of staining positivity, further highlighting the need for a standardized scoring methodology for Ki-67. The challenge of reporting Ki-67 accurately and consistently led us to assess the intra-laboratory interobserver variability in the estimation of Ki-67 PI using the method suggested by the IBCWG. This is a retrospective study of cases where estrogen receptor (ER) positive, HER-2/neu negative invasive breast carcinomas had also been evaluated for Oncotype DX™. The Ki-67 PI was correlated with the recurrence score (RS) obtained from Oncotype DX.

MATERIALS AND METHODS

Using the laboratory information system to perform a search, 27 consecutive cases of invasive breast carcinomas between 2010 and 2012 where Oncotype DX assay was also performed were included in the study. ER, progesterone receptor (PR), and HER-2/neu data was...
obtained from pathology reports. IHC for ER, PR and HER-2/neu was performed using the SP1, 1E2 and 4B5 rabbit monoclonal antibodies, respectively, and iVIEW detection on the benchmark XT (Ventana, Tucson, AZ, USA). At our institution, the ER and PR results are reported based on the percentage of positive cells showing non, weak, moderate or strong staining. HER-2/neu expression is documented as per the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) guidelines.[3] Only the breast carcinomas with positive expression of ER and negative expression for HER-2/neu protein (IHC and/or fluorescence in situ hybridization) were included in the study. Representative hematoxylin and eosin (H and E) stained slides were retrieved from the files of the pathology department. All sections had been fixed in 10% neutral phosphate buffered formalin for duration varying from 8 to 48 h, according to ASCO/CAP guidelines.[3] The same paraffin block was selected that had been previously used for Oncotype DX testing (Genomic Health, Redwood city, CA, USA). Immunostaining for Ki-67 (Rabbit monoclonal antibody, 30-9 clone, pre-diluted, benchmark XT, Ventana, Tucson, AZ, USA) was performed on a 5 µm section of paraffin embedded tissue. The immunostain for Ki-67 and the corresponding H and E slides were independently reviewed by three pathologists (VM, XZ, RT). Relevant clinical and pathology data including tumor size, histological type, Nottingham grade and hormone receptor status were retrieved from the pathology reports. The Oncotype DX reports were reviewed and the RS was recorded.

The recommendations made by the IBCWG for scoring the Ki-67 staining were discussed by the three pathologists and an agreement was reached on the counting methods. Only tumor cells with nuclear staining, regardless of intensity, were counted and reported as a percentage of positively staining cells among the total number of tumor cells.[3] If the staining was homogenous [Figures 1a and 2a] a minimum of three randomly selected high power fields (>40) were counted; if there was a gradient of increasing staining toward the tumor’s invasive edge [Figures 1b and 2b] three fields at the periphery were scored; if hot spots (areas where Ki-67 staining nuclei was particularly dense) were present in otherwise homogenous staining [Figures 1c and 2c], at least one hot spot was included in the count.[3] The three patterns of staining described above are referred to as patterns A, B and C, respectively. A minimum of 500 cells were counted irrespective of the number of fields. There is no current consensus recommendation for cut-off to separate “low Ki-67” from “high Ki-67.” Most studies in the literature have used staining levels of 10-20% to dichotomize populations; results in our study were evaluated in increasing intervals of 5% points.[2]

The inter-rater agreement was calculated using intra-class correlation coefficient (ICC). The mean Ki-67 PI was calculated for the three observers and the value was compared with the RS in 26 cases. One case did not have adequate material on the paraffin block and the Oncotype DX test could not be performed. As per the literature, the RS classifies tumors into three risk categories: Low (0-17), intermediate (18-30) and high (>31).[3] Pearson correlation (r) was calculated between Ki-67 and RS.

**RESULTS**

There were 27 cases in the study. The mean patient age was 58 years (range: 44-79 years), and they were all female. 16 specimens were lumpectomies and 11 specimens were simple mastectomies. Sentinel lymph node dissection was performed in 26 cases; in 5 cases, micro-metastasis was present in at least 1 lymph node. Of the 27 cases, there were 22 cases of invasive ductal carcinoma (10 cases of Nottingham Grade 1; 11 cases of Nottingham Grade 2; 1 case of Nottingham Grade 3) and 5 cases of invasive lobular carcinoma (4 cases of classic type, Nottingham Grade 1; 1 case of pleomorphic type, Nottingham Grade 3). Accompanying ductal carcinoma in situ and/or lobular carcinoma in situ was identified in 23 cases. The mean tumor size was 2.1 cm (range: 0.5-4.0). All tumors showed moderate to strong positive expression of ER (80-100% tumor cells) [Table 1]. The IHC expression of PR was strongly positive in 18 cases, weakly positive in 3 cases and negative in 6 cases [Table 1]. The results of ER and PR

| Table 1: Number of cases (n) in different risk categories based on tumor grade, PR expression, Ki-67 index |
|-----------------|-----------------|-----------------|-----------------|
| **Prognostic parameters** | **Low risk (RS<18) (n)** | **Intermediate risk (RS 18-30) (n)** | **High risk (RS>31) (n)** |
| **Grade**       |                  |                  |                  |
| 1               | 4                | 10               | 1                |
| 2               | 6                | 2                | 2                |
| 3               | 0                | 1                | 3                |
| **PR (%)**      |                  |                  |                  |
| Negative        | 0                | 4                | 2                |
| <5              | 0                | 1                | 0                |
| 5-25            | 0                | 1                | 0                |
| 26-50           | 0                | 1                | 0                |
| 51-75           | 0                | 1                | 0                |
| 76-100          | 10               | 5                | 0                |
| **ER (%)**      |                  |                  |                  |
| 80-100          | 10               | 13               | 3                |
| **Ki-67 (%)**   |                  |                  |                  |
| 0-5             | 2                | 0                | 0                |
| 6-10            | 0                | 5                | 0                |
| 11-15           | 1                | 0                | 0                |
| 16-20           | 4                | 1                | 0                |
| 21-25           | 2                | 3                | 0                |
| 26-30           | 0                | 2                | 0                |
| 31-35           | 1                | 1                | 0                |
| >36             | 0                | 1                | 3                |

PR: Progesterone receptor, RS: Recurrence score, ER: Estrogen receptor
immunoexpression were similar to the ER and PR score obtained from the Oncotype DX assay.

Scoring of Ki-67 PI was performed on all 27 cases independently by 3 different pathologists. At least 2 of 3 pathologists agreed in interpretation of pattern A in 18 cases, pattern C in 6 cases and pattern B in 2 cases. All three pathologists disagreed on the staining pattern in only 1 case.

In the assessment of Ki-67 PI, the three observers were overall 89.1% in agreement (ICC = 0.891, 95% confidence interval [CI]: 0.806-0.945) [Figure 3]. The agreement under different risk categories, was good in the low risk group (ICC = 0.7945, 95% CI: 0.5272-0.9382) and intermediate risk groups (ICC = 0.8300, 95% CI: 0.6342-0.9389). The agreement was fair in the high risk group (ICC = 0.6374, 95% CI: 0.0501-0.9877).

RS (Oncotype DX) was analyzed in 26 cases; there were 10 cases with low risk (mean group RS: 10.1; mean Ki-67: 17.4; Ki-67 range: 3-33%); 13 with intermediate risk (mean group RS: 21.4; mean Ki-67: 19.5; Ki-67 range: 6-43%) and 3 with high risk (mean group RS: 53; mean Ki-67: 70; Ki-67 range: 55-91%). An overall very strong positive correlation between Ki-67 and Oncotype DX ($r = 0.78464, P < 0.0001$) [Figure 4] was obtained. However, when analyzed under separate risk groups, the very strong positive correlation between Ki-67 and Oncotype DX was limited to only high risk groups ($r = 0.96077, P = 0.1789$); all three cases in the high risk group had a Ki-67 > 30% [Table 1]. The unexpected high Ki-67 score in the low risk category (Ki-67 > 10% in 8 of 10 cases including three cases with a score > 20%) [Table 1], resulted in a strong negative correlation in low risk group (-0.46773, $P = 0.1728$) [Figure 5]. A weak positive correlation was obtained in the intermediate risk group ($r = 0.2075, P = 0.4964$).

When RS was compared with other available parameters, a strong and significant negative correlation between RS and PR expression ($r = -0.63804, P < 0.0005$) was identified. The 10 low risk category cases were strongly positive for PR; the 3 high risk category cases showed weak to negative expression for PR; the intermediate risk category cases had PR positivity ranging from negative to strongly positive [Table 1].

There was no correlation of RS or Ki-67 with size of the tumor and grade of the tumor [Table 2].

**DISCUSSION**

Ki-67 PI is an established and a useful marker of cell proliferation and has been offered as a prognostic marker in early breast cancer. However, the predictive value of a high Ki-67 labeling index for response to adjuvant chemotherapy is unclear. Several studies have reported that among patients with node-negative, hormone receptor-positive breast cancer, a high Ki-67 PI has worse disease free survival trend than a low Ki-67 PI but the Ki-67 PI does not predict the efficacy of adjuvant chemotherapy.$^{[1,7]}$

Use of proliferation markers such as Ki-67 PI, along with multi-gene assays has been recommended to determine the benefit of chemotherapy in addition to hormone therapy in node negative ER positive breast cancer.$^{[6]}$ The Oncotype DX gene test is a commercially available reverse transcriptase polymerase chain reaction (RT-PCR) assay of 21 genes which uses a specific algorithm to calculate the RS for ER positive breast cancers. In the multi-gene assay, five genes including Ki-67 reflect the proliferative status of the tumor.$^{[6,8]}$ However, the high cost of this assay limits its use and has led to evaluation of various alternative histopathological and clinical parameters as predictive factors.$^{[9]}$ Due to its simplicity, IHC for Ki-67 is one of the widely investigated biological markers of breast cancer.$^{[9]}$ Despite consistent data on Ki-67 PI as a prognostic marker in early breast cancer, its role in breast cancer management remains uncertain. The

![Figure 1](image1.png)

**Figure 1:** Illustration of three high power fields (red circles) selected based on pattern of staining: (a) homogenous, (b) gradient of increasing staining toward the tumors invasive edge and (c) hot spots

![Figure 2](image2.png)

**Figure 2:** Breast carcinoma with different patterns of Ki-67 immunostaining ($\times200$); (a) homogenous, (b) gradient of increasing staining toward the tumors invasive edge and (c) hot spots
The reproducibility of Ki-67 measurement by IHC was the current method of choice for measuring and monitoring tumor proliferation in standard pathology specimens. They also acknowledged poor agreement on the precise clinical usefulness of Ki-67, substantial heterogeneity and variable levels of validity in methods of assessment. To address this issue they proposed guidelines on the preferred pre-analytical, analytical, staining and scoring methodologies for its analysis. In a much recently published study, the Working Group studied intra- and inter-laboratory reproducibility of IHC assays for Ki-67 in breast cancer among a group of highly experienced pathology laboratories. They found that intra-laboratory reproducibility was high (ICC = 0.94; 95% CI = 0.93-0.97); however the inter-laboratory reproducibility was only moderate (ICC = 0.59-0.71, 95% CI = 0.37-0.78). In the present study, we assessed the intra-laboratory reproducibility of Ki-67 index in compliance with the guidelines. Using the IBCWG recommendations the three pathologists were able to achieve an agreement of 89.1% in the assessment of Ki-67. The agreement was good in low and intermediate risk groups but was only fair in high risk groups. Greater interobserver variability among tumors with higher labeling indices as compared with tumors with labeling indices closer to 0 has been reported in brain tumors. This variability may be attributed to the differing lower thresholds for interpreting positivity. The reported association of specific cut-off values with a higher level of interobserver variability also raises the possibility of grading the Ki-67 index subjectively as low, moderate, or high, without numeric cut-offs. Although, our study is limited by a small study group, we found that the high risk category is more likely to be associated with a very high Ki-67 (>30%). Similar observations have been made by other authors. Even though a high Ki-67 PI confers a worse prognosis in early breast cancer, it does not have an independent prognostic value and cannot be the sole determinant to identify the subgroup. Lack of agreement among researchers about the usefulness of Ki-67 as a prognostic index and the quest to improve on traditional predictive and prognostic factors has persisted and resulted in the emergence of numerous multi-gene expression profiles. One such assay is Oncotype DX, a commercially available RT-PCR-based assay that is used to predict the risk of RS and chemotherapy benefit in patients with ER positive, node negative breast cancers. The RS has been shown to have a strong correlation with tumor grade, PR levels and Ki-67 in subsets of ER positive patients. In our study, we did not find any correlation between RS and grade, but there was a significant negative correlation with PR expression \( r = -0.63 \) and a positive correlation \( r = 0.78464, \ P \leq 0.0001 \) between Ki-67 and RS.
RS had a strong positive PR expression and all the 3 cases with high RS had lacked PR expression. Association of lower expression of PR with higher Oncotype RS has been reported previously by other authors.\textsuperscript{5,9,18,19} The Oncotype DX RS relies heavily on parameters already available from routine pathologic examination, and consideration of progesterone receptor status may aid in selection of patients most likely to benefit from ancillary testing.\textsuperscript{19} The positive correlation between Ki-67 and RS is compromised by the presence of outliers, as evident in the scatter plot [Figure 4]. The wide range (3-33%) and high Ki-67 values in the low risk category is a limiting factor in establishing a cut-off for Ki-67 to use as a prognostic/predictive index. The strong correlation between Ki-67 PI and RS established by other studies is also limited by presence of outliers (unexpected high Ki-67 value in low risk groups) and by the small size of their study groups.\textsuperscript{5,20,21}

Though limited by a small number of cases our study suggests that use of the recommended guidelines by the IBCWG may help overcome the challenge of a lack of a standardized scoring system to calculate Ki-67 PI. However, the wide range and unexpected high Ki-67 in some patients with low RS is a limiting factor in establishing a cut-off. The role of a weak or negative progesterone expression needs to be further investigated in a larger cohort of high risk ER positive patients.

\section*{Acknowledgment}

Would like to acknowledge the contribution of Marsha McCray, Office Manager, Surgical Pathology for retrieving the slides and procuring the Oncotype Dx reports on all the cases.

\section*{References}


