Standardized Assessment of Ki-67 in Breast Cancer Patients Using Virtual Slides and an Automated Analyzer in Comparison to Central/Local Pathological Assessments

Yoshio Mizuno1*, Hiromi Fuchikami1, Tsuneo Natori2, Naoko Takeda1,3, Yuko Inoue3, Junichi Yamada4, Hiroaki Abe4, Hiroshi Seto5, Kazuhiko Sato1

1Department of Breast Oncology, Tokyo-West Tokushukai Hospital, Tokyo, Japan; 2Clinical Testing Business, Testing Division, SRL, Inc., Tokyo, Japan; 3Inoue Ladies Clinic, Tokyo, Japan; 4Department of Clinical Pathology, Tokyo-West Tokushukai Hospital, Tokyo, Japan; 5Seto Hospital, Tokorozawa, Japan.

Received December 30th, 2013; revised January 18th, 2014; accepted January 25th, 2014

Copyright © 2014 Yoshio Mizuno et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for SCIRP and the owner of the intellectual property Yoshio Mizuno et al. All Copyright © 2014 are guarded by law and by SCIRP as a guardian.

ABSTRACT

Purpose: To standardize the methods to measure Ki-67, there is an interest in automating the assessment of Ki-67. Therefore, we reviewed the possibility of introducing an automated analyzer to standardize the Ki-67 evaluation method. Methods: We retrospectively reviewed a clinical database of patients who underwent surgery for early breast cancer at Tokyo-West Tokushukai Hospital. Among them, those who underwent preoperative core needle biopsy (CNB) were enrolled. The concordance rates of estrogen receptor (ER), progesterone receptor (PgR), human epidermal growth factor receptor 2 (HER2), and Ki-67 by local pathologists were reviewed (valuations made by local pathologists), and nonmatching cases (from August 2008 to October 2011) were reassessed both by central review and using an automated analyzer with virtual slides. The results were compared with the evaluations made by local pathologists, and we reexamined the concordance rate by using central review and the automated analyzer. Results: The concordance rate of Ki-67 evaluations made by local pathologists in the preoperative CNB and surgical specimens was 78.7% in 287 cases pathologically assessed from October 2008 to March 2013. This rate was significantly lower (p < 0.01) than that of ER (95.6%), PgR (88.5%), and HER2 (91.6%). Reassessment of the 37 cases of nonmatching Ki-67 values from 2008 to October 2011 using central review and an automated analyzer resulted in clear improvement in matching of 22 (92.1%) and 24 (93.1%) of 37 cases, respectively. Conclusion: The concordance rate of Ki-67 in preoperative CNB and surgical specimens was lower than that of other biological markers; however, they were nearly equal by reassessment using central review and an automated analyzer.

KEYWORDS

Ki-67; Core Needle Biopsy; Concordance Rate; Virtual Slides; Automated Analyzer; Standardized Assessment

1. Introduction

Expression of the nuclear nonhistone protein Ki-67 (proliferation-related Ki-67 antigen) is associated with cell proliferation during interphase and commonly used as a predictive and prognostic marker of breast cancer [1-7]. In particular, in neoadjuvant endocrine therapy, differences in pre- and post-treatment using Ki-67 are reportedly related to preoperative hormone therapy as a predictor of prognosis and treatment outcome [8-12]. However, there are certain limitations to the Ki-67 index, including standardization of a measurement method, reproducibility of measurement results, and establishment of cut-off values [13-16]. The core needle biopsy (CNB) procedure is almost as accurate as immunohistochemical
analysis of surgical specimens for breast cancer diagnosis and is now widely accepted as the standard diagnostic procedure [17-20]. However, previous reports have shown variations in Ki-67 expression among pretreatment CNB specimens and post-treatment surgical specimens [21-23]. Therefore, it is important to consider potential discrepancies in Ki-67 values between CNB and surgical specimens, even in patients who receive no preoperative treatment, to standardize Ki-67 measurements in preoperative hormone therapy.

In preanalytic validity, formalin fixation condition can adversely affect the measurement results. Equally in analytical validity, tumor heterogeneity and the lack of standard measuring method by pathologists may affect Ki67 measurement [11]. We had previously reported that these factors might affect the concordance rate for Ki-67 expression between CNB and surgical specimens, and we suggest that standardization of the measurement method by pathologists is most important [21]. However human observers spend considerable time and effort assessing large tissue areas and performing central reviews in general medical facilities which is difficult. Consequently, this retrospective study was conducted to assess the possibility of introducing an automated analyzer for assessment by comparing diagnoses made by local pathologists and reassessment by central review to standardize an evaluation method of the Ki-67 LI in primary breast cancer.

2. Methods

2.1. Patients and Materials

We retrospectively analyzed the medical records of 446 primary breast cancer patients who underwent their first surgeries from August 2008 to March 2013 and selected 287 who underwent preoperative CNB at Tokyo-West Tokushukai Hospital (Tokyo, Japan). For CNB, a 16- or 18-gauge automated needle device with a 22-mm throw biopsy gun was used. Three or more CNB specimens were obtained from each patient and placed in 20% formalin for 6-48 h. The original tumors were fixed in buffered formalin and embedded in paraffin. One representative tissue block for each tumor was selected for routine immunohistochemical (IHC) analysis of ER, PgR, HER2, and Ki-67 levels, which were conducted by many unspecified registered local pathologists. Patients with ductal carcinoma in situ and those that received neoadjuvant chemotherapy were excluded.

2.2. Ki-67 Scoring

Immunohistochemical staining was quantitatively evaluated by light microscopy, in which the entire tissue section was scanned at low-power magnification (10×) to determine areas with the highest number of positive nuclei (hot-spots) within the invasive component. Ki-67 was expressed as the percentage of cells positive for mindbomb E3 ubiquitin protein ligase 1 (MIB-1) among a total of at least 1000 malignant cells at high-power magnification (40×). Nuclear staining of the tumor cells was considered negative if 14% or fewer of the cells were stained for Ki-67 and as positive if more than 14% were stained for Ki-67. An MIB-1 clone (Dako, Carpinteria, CA, USA) was used for immunohistochemical analysis of Ki-67.

2.3. Assessment of Other Biomarkers

An IHC score of at least 10% positive cells was used to define ER/PgR positivity. A positive score for HER2 was either HER2 3+ by IHC analysis (defined as uniform intense membrane staining of >30% of invasive tumor cells) or fluorescence in situ hybridization (ratio of HER2 to chromosome 17 centromere of >2.0).

2.4. Reassessment of Ki-67 by Central Review

A central review by a professionally trained physician was performed by scanning magnification to count at least 1000 cells in the most densely labeled areas. For all nonmatching cases, the percentage of tumor cells with any nuclear staining was recorded. The central review used calculations based on the hot-spot counting method because counting the area with the highest number of positive cells was more reproducible compared to random counting.

2.5. Virtual Slides and Automated Analyzer

The 37 cases with nonmatching Ki-67 scores (74 slides) were scanned using the slide scanner with a 20× objective and reassessed using the image management software system. The image management software system employs image analysis techniques with predefined parameters to obtain Ki-67 scores. The hot-spot counting method was used with the automated analyzer. Nuclear identification was automatically performed using image analysis algorithms, which involved the following steps: 1) image enhancement, in which the image contrast is adjusted to make it suitable for analysis; 2) identification of the epithelial area, which is defined as the image region in which there was a possibility of the presence of epithelial cells; 3) identification of the nucleus; 4) classification of cells based on the extent and intensity of nuclear staining; and 5) computing the score. The algorithm will reject elongated nuclei regardless of the overall cell shape; therefore, tumors containing large numbers of
cells with elongated nuclei must be manually evaluated.

2.6. Comparison of Automated and Central/Local Pathology Assessment

All cases were evaluated by local registered pathologists. Although many different systems for grading pathological responses have been proposed, no standard method has yet been adopted. The concordance rates for assessment of ER, PgR, HER2, and Ki-67 by local pathologists were reviewed, and in cases of nonmatching Ki-67 scores, the tumor diameters (approximately indicative of tumor heterogeneity) and surgical method (approximately indicative of formalin fixation condition) were evaluated.

Further, the 37 cases with nonmatching Ki-67 scores (74 slides) from August 2008 to October 2011 were re-assessed by central review using an automated analyzer with virtual slides and compared with pathological evaluations. Reassessments of the 37 cases with nonmatching Ki-67 values, as determined by central review and the automated analyzer, were calculated using a hot-spot counting method because areas with the highest number of positive cells were used rather than random counting to enhance reproducibility.

Statistically significant differences between the concordance rates of the two specimen types were evaluated using the Wilcoxon t-test. To evaluate the consequence of formalin and genetic heterogeneity, parameters, such as the surgical method and tumor size, were analyzed using the chi-squared test.

According to the policies of our institutional ethics committee, general consent was obtained from all patients receiving medical care.

3. Results

The mean patient age was 56.4 years (median, 55.5 years; range, 30 - 91 years). Seventy-three patients ultimately underwent mastectomy and the remainder underwent breast-conserving surgery. A total of 227 cases were ER-positive and 47 ER-negative for both CNB and surgical specimens. Six cases had ER-positive CNB specimens and negative surgical specimens, and seven cases had negative CNB specimens and positive surgical specimens. A total of 163 cases had PgR-positive CNB and surgical specimens, 91 had negative CNB and surgical specimens, 16 had positive CNB specimens and negative surgical specimens, and 17 had negative CNB specimens and positive surgical specimens. Regarding HER2 expression, 40 cases had positive CNB and surgical specimens, 210 had negative CNB and surgical specimens, 16 had positive CNB specimens and negative surgical specimens, and seven had negative CNB specimens and positive surgical specimens. Regarding Ki-67, 144 cases had positive CNB and surgical specimens, 82 had negative CNB and surgical specimens, 37 had positive CNB specimens and negative surgical specimens, and 24 had negative CNB specimens and positive surgical specimens.

The concordance rates of marker expression between the CNB and surgical specimens are shown in Tables 1-4. In our series, the concordance rate for Ki-67 expression between the two specimen types was 78.7%, which was significantly lower than that for ER, PgR, and HER2 (95.5%, 88.5%, and 91.6%, respectively; Table 5).

Analytical results of tumor staging (approximately indicative of tumor heterogeneity) and the surgical methods (approximately indicative of differences in formalin fixation conditions) are shown in Table 6. No significant difference in parameters, such as tumor stage (pT1 vs. ≥pT2) and surgical method (mastectomy vs. breast-conserving surgery), were observed between the two patient groups.

### Table 1. Concordance rates between CNB<sup>a</sup> and surgical specimens for ER<sup>b</sup> status.

<table>
<thead>
<tr>
<th></th>
<th>Positive surgical specimens</th>
<th>Negative surgical specimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (CNB)</td>
<td>227</td>
<td>6</td>
</tr>
<tr>
<td>Negative (CNB)</td>
<td>7</td>
<td>47</td>
</tr>
</tbody>
</table>

Concordance rate for ER: 95.5%; <sup>a</sup>core needle biopsy; <sup>b</sup>estrogen receptor.

### Table 2. Concordance rates between CNB<sup>a</sup> and surgical specimens for PgR<sup>b</sup> status.

<table>
<thead>
<tr>
<th></th>
<th>Positive surgical specimens</th>
<th>Negative surgical specimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (CNB)</td>
<td>163</td>
<td>16</td>
</tr>
<tr>
<td>Negative (CNB)</td>
<td>17</td>
<td>91</td>
</tr>
</tbody>
</table>

Concordance rate for PgR: 88.5%; <sup>a</sup>core needle biopsy; <sup>b</sup>progesterone receptor.

### Table 3. Concordance rates between CNB<sup>a</sup> and surgical specimens for HER2<sup>b</sup> status.

<table>
<thead>
<tr>
<th></th>
<th>Positive surgical specimens</th>
<th>Negative surgical specimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (CNB)</td>
<td>40</td>
<td>16</td>
</tr>
<tr>
<td>Negative (CNB)</td>
<td>7</td>
<td>210</td>
</tr>
</tbody>
</table>

Concordance rate for HER2: 91.6%; <sup>a</sup>core needle biopsy; <sup>b</sup>human epidermal growth factor receptor 2.

### Table 4. Concordance rates between CNB<sup>a</sup> and surgical specimens for Ki-67 expression.

<table>
<thead>
<tr>
<th></th>
<th>Positive surgical specimens</th>
<th>Negative surgical specimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (CNB)</td>
<td>144</td>
<td>37</td>
</tr>
<tr>
<td>Negative (CNB)</td>
<td>24</td>
<td>82</td>
</tr>
</tbody>
</table>

Concordance rate for Ki-67: 78.7%; <sup>a</sup>core needle biopsy.
Table 5. Comparison of concordance rates between ER, PgR, and HER2 with that of Ki-67.

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Concordance rates</th>
<th>Comparison of concordance rates with Ki-67</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>95.5%</td>
<td><em>p</em> &lt; 0.01</td>
</tr>
<tr>
<td>PgR</td>
<td>88.5%</td>
<td><em>p</em> &lt; 0.01</td>
</tr>
<tr>
<td>HER2</td>
<td>91.6%</td>
<td><em>p</em> &lt; 0.01</td>
</tr>
<tr>
<td>Ki-67</td>
<td>78.7%</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. No significant difference between tumor stage and surgical methods.

<table>
<thead>
<tr>
<th>Type</th>
<th>Match (n = 226)</th>
<th>Non-match (n = 61)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>56.1 (30 - 91)</td>
<td>57.8 (35 - 79)</td>
<td>= n.s</td>
</tr>
<tr>
<td>Tumor stage</td>
<td>pT1</td>
<td>144 (63.7%)</td>
<td>38 (62.3%)</td>
</tr>
<tr>
<td></td>
<td>≥pT2</td>
<td>82 (36.3%)</td>
<td>23 (27.7%)</td>
</tr>
<tr>
<td>Types of surgery</td>
<td>BCS</td>
<td>144 (63.7%)</td>
<td>36 (59.0%)</td>
</tr>
<tr>
<td></td>
<td>Mastectomy</td>
<td>82 (36.3%)</td>
<td>25 (41.0%)</td>
</tr>
</tbody>
</table>

Reassessment (October 2008-October 2011) of the 37 cases with unequal Ki-67 scores after assessment by local pathologists resulted in matching of 22 cases by central review and 24 using the automated analyzer. The concordance rate of the re-examined specimens among the Ki-67 discordant group by central review and using the automated analyzer improved to 92.1% and 93.1%, respectively, but this was not significantly different from that of the other receptors.

4. Discussion

The results of the present study showed that the concordance rate assessed by local pathologists between Ki-67 expression in the preoperative CNB and surgical specimens was 78.7% in the 287 cases from October 2008 to March 2013, which was significantly lower (p < 0.01) than that for ER (95.5%), PgR (88.5%), and HER2 (91.6%). Among the nonmatching cases, no differences were observed in tumor diameter or surgical method. Reassessment (October 2008-October 2011) of the 37 cases not matching for Ki-67 after local pathologist assessments resulted in matching in 22 cases by central review and 24 cases using the automated analyzer. The concordance rate of the re-examined specimens among the Ki-67 discordant group by central review and using the automated analyzer improved to 92.1% and 93.1%, respectively, which was not significantly different from that of the other receptors.

The International Ki67 in Breast Cancer Working Group published their recommendations for Ki67 assessment in breast cancer in 2011. However, these recommendations included no established quality assurance schemes to ensure that the procedures for Ki67 analysis in one laboratory lead to comparable scores in others [11]. Thus, the direct application of specific cut-off rates for comparison must be considered unreliable unless the analyses were performed in a high-volume laboratory with its own reference data. However, this article did not refer to the use of an automated analyzer among its recommendations.

Some studies have analyzed the reliability of Ki-67 assessment in breast cancer using the automated analyzer in comparison to human counts. Fasanella et al. [24] analyzed 315 consecutive breast cancer specimens immunostained for Ki-67 that were examined both by an experienced pathologist and computer-assisted image analysis (CAIA) and showed a correlation between the human and CAIA evaluations, although the CAIA values were slightly lower [24]. Mohammed et al. [25] compared visual and automated Ki67 assessment methods and survival in 379 breast cancer patients and reported that the methods were in excellent agreement. Furthermore, univariate analysis revealed that visual and automated Ki67 assessment methods were associated with overall cancer-specific survival in patients with invasive ductal breast cancer and in those who received endocrine therapy.

In this study, we analyzed the concordance rates between tumor diameter and its equivalence to tumor heterogeneity as well as compared Ki-67 expression to determine the effect on tumor heterogeneity, but found no significant correlation. Furthermore, no significant correlation was observed among the concordance rates between the surgical method and its equivalence to formalin fixation conditions as well as Ki-67 regarding the effect on formalin fixation. To assess the lack of a standard measurement method among pathologists, we re-examined cases showing a discordance in Ki-67 assessment through a central review. The concordance rate between preoperative CNB specimens and surgical specimens assessed by local pathologists was lower for Ki-67 expression than that of the other biological markers. However, these rates became nearly equal after reassessment of Ki-67 by central review and using the automated analyzer. The automated analyzer has the advantage of measuring a much greater number of cells and reduces the time to acquire results compared to human observations. Manual observation requires considerable time and effort to assess large tissue areas and performing central reviews in general medical facilities is difficult. Our results showed that the use of an automated analyzer for Ki-67 assessment can solve this problem.

Major limitations to this study included the retrospective nature of the review with limited objective endpoints.
and case selection was non-randomized and limited to a single oncology center. Furthermore, reassessment by central review and using an automated analyzer was performed for only the nonmatching cases. Despite these limitations, this retrospective analysis highlighted the importance for standardization of pathological methods for assessment of Ki-67 in clinical use.

5. Conclusion

In conclusion, central review and the use of an automated analyzer can improve the accuracy of Ki67 assessment. The results of this study confirmed the necessity of a standardized evaluation method for Ki-67 expression in breast cancer to overcome the disadvantages of variable counting methods and measurement sites. Our data suggest that the use of an automated analyzer can assist in the standardization of a Ki-67 evaluation method.

Competing interests

The authors declare no competing interests.

Acknowledgments

The author would like to thank Enago (www.enago.jp) for the English language review. This study was presented in part at the 35th Annual San Antonio Breast Cancer Symposium (poster discussion).

Disclosure

The authors declare no conflicts of interest.

REFERENCES


http://dx.doi.org/10.1136/jclinpath-2012-201085


