Disclosure Information

I hereby declare that I have had business or personal interests in the following industrial enterprises since 1 September 2016:

<table>
<thead>
<tr>
<th>Name of the enterprise / Nature of the interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nothing to declare</td>
</tr>
</tbody>
</table>
Do snapshot and whole slide image analysis display with differences of Ki-67 stained invasive breast carcinoma specimens?

Dr. Essam E. Ayad
Prof. of Pathology & Head of Digital Pathology Unit, Faculty of Medicine, Cairo University
Consultant Pathologist, Italian Hospital In Cairo
Introduction

- Breast cancer is the second most common cancer in the world and, by far, the most frequent cancer among women.

- The determination of the growth fraction with the antibody Ki-67 has been widely used in histopathology.

- In general, the Ki-67 score is defined as the percentage of total number of tumour cells with nuclear staining.
ROLE OF KI 67 IN BREAST CANCER:

- Distinguishing between luminal A and Luminal B molecular subgroups (St Gallen Consensus Meeting In 2011)

- Selection of cases for Chemotherapy: Patients with high Ki 67 index are the tumor subgroup that should receive chemotherapy and hormone therapy

- Ki 67 as prognostic and predictive factor: high Ki 67 index value was associated with poor prognostic clinicopathological factors, high risk of relapse, significantly shorter Disease Free Interval and a significantly lower overall survival rates
However, this marker has not yet been implemented for routine clinical use due to the lack of standardized methodologies for measurement:

1. Which part of the field to assess: (whole slide examination, hotsort,)

2. The methods of quantification of positive nuclei: eyeballing or counting several hundred consecutive nuclei

3. Widely varying cutoff values to classify patients into “Ki67 high” or “Ki67 low” risk groups (2011 St. Gallen International Consensus Meeting Conference Panel recommended a cutoff of 14% but 2 years later, it was upgraded to 20% and still no standard cut off point is universally admitted)
Recently, Digital pathology technologies, enabling high-resolution scanning of microscopy slides, brings great efficiencies in data storage, transfer and usage in research, clinical practice and education.

Development of computer-aided image analysis software could help to eliminate the inherent variability of pathologist-based scoring limiting the use of multiple immunohistochemistry markers in the daily routine.
AIM OF THE WORK

- Present and validate an easy-to-use, standardized and accurate Ki-67 scoring method in breast cancer by:
  
  • Comparing inter-observer agreement of Ki-67 index assessment using optical and digital microscopy.
  
  • Comparing the results of manual assessment of Ki-67 index versus those of the automated quantitative assessment of both snap shots & the WSI of the studied cases.
MATERIAL AND METHODS

Study was done on histological sections obtained from paraffin blocks of 100 lumpectomy/excisional/core biopsy specimens from 100 patients diagnosed with invasive breast cancer.

- The 100 paraffin blocks were sectioned and stained with Hematoxylin and Eosin (H&E) for reevaluation followed by staining with immunohistochemical Ki 67 stain.

- Slides were then scanned by iScan.

- Stained glass slides as well as virtual slides were examined by 3 different pathologists.
MANUAL ASSESSMENT OF KI 67 INDEX:

- The 3 pathologists were asked to estimate the Ki-67 index for each of the 100 cases by examining the whole slide on optical microscopy and on computer screen and provide a score in the range of 0 to 100 corresponding to the percentage of positive tumors cells: **Continuous score**

- The results were then classified into 2 categories using 2 different cutoff points: **Categorical scores:**
  - **First:** Cutoff point = 14%: ‘High Ki 67’ status if the ki 67 index $\geq 14 \%$ and ‘low Ki 67’ status otherwise.
  - **Second:** Cutoff point = 20%: ‘High Ki 67’ status if the ki 67 index $\geq 20 \%$ and ‘low Ki 67’ status otherwise.
QUANTITATIVE DIGITAL ANALYSIS OF KI67

Whole scanned slide was examined then; multiple snapshots were captured with a (x40) objective covering almost the whole scanned slide (15-50 snapshots per case).

A digital quantitative analysis of Ki 67 proliferative index was performed using PathIA program for all the snap shots taken from all cases.

A digital quantitative analysis of Ki 67 proliferative index was performed using Halo software for the WSI.
Results

Low Ki 67 Index < 14%
High Ki 67 index
<p><strong>RESULTS:</strong></p>

- **Manual assessment of Ki 67 by 3 observers on optical microscopy (Continuous score):**

<table>
<thead>
<tr>
<th>Observers</th>
<th>Mean ± SD (N=100)</th>
<th>Range (N=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer A</td>
<td>30.22±20.8</td>
<td>5%-90%</td>
</tr>
<tr>
<td>Observer B</td>
<td>29.62±21.16</td>
<td>1%-85%</td>
</tr>
<tr>
<td>Observer C</td>
<td>32.43±22.38</td>
<td>1%-85%</td>
</tr>
</tbody>
</table>

- **Manual assessment of Ki 67 by 3 observers on virtual slides**

<table>
<thead>
<tr>
<th>Observers</th>
<th>Mean ± SD (N=100)</th>
<th>Range (N=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer A</td>
<td>31.93±24.62</td>
<td>1%-95%</td>
</tr>
<tr>
<td>Observer B</td>
<td>35.79±23.13</td>
<td>2%-90%</td>
</tr>
<tr>
<td>Observer C</td>
<td>21.61±16.72</td>
<td>0%-80%</td>
</tr>
</tbody>
</table>
RESULTS:

- **Continuous score: manual assessment**

**Optical microscopy**

- Observer A vs. Observer B
- Observer A vs. Observer C
- Observer B vs. Observer C

**Virtual slides**

- Observer A vs. Observer B
- Observer A vs. Observer C
- Observer B vs. Observer C

- **overall Inter-observer agreement on optical microscopy:** almost perfect: (CC) ranging from 0.847 to 0.937 (p value <0.01)

- **slightly better than inter-observer agreement on virtual slides** which was substantial to almost perfect : (CC) ranging from 0.717 to 0.882 (p value <0.01).
* Quantitative assessment of Ki67 using continuous and categorical scores:

<table>
<thead>
<tr>
<th>Quantitative assessment of Ki67</th>
<th>Continuous Score</th>
<th>Categorical Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cut off 14%</td>
</tr>
<tr>
<td></td>
<td>Mean ± SD (N=100)</td>
<td>Range (N=100)</td>
</tr>
<tr>
<td></td>
<td>39.26±27.43</td>
<td>0.5-95.39</td>
</tr>
</tbody>
</table>
The overall agreement between the manual and automated evaluation of Ki 67 was substantial (CC ranging from 0.696 and 0.791 on optical microscopy and 0.662 and 0.792 on computer screen (p value <0.01)).

Correlation between optical microscopy results provided by the 3 observers and quantitative analysis results was slightly better than the correlation between virtual slides results and automated assessment results.
Categorical score: manual assessment: cut off point 14%

**Optical microscopy:** Inter-observer agreement: substantial to almost perfect: kappa values ranging from 0.696 and 0.925 (p value<0.01).

![Bar charts showing optical microscopy results](chart1.png)

**Virtual slides:** Inter-observer agreement: Moderate to substantial: kappa values ranging from 0.527 and 0.666 (p value<0.01).

![Bar charts showing virtual slides results](chart2.png)

Inter-observer agreement was better on optical microscopy while using 14% as cut off.
Categorical score: manual assessment: cut off point 20%

**Optical microscopy:** Inter-observer agreement: substantial to almost perfect: kappa values ranging from 0.672 and 0.938 (p value<0.01).

**Virtual slides:** Inter-observer agreement: Moderate to substantial: kappa values ranging from 0.525 and 0.612 (p value<0.01).

Inter-observer agreement was better on optical microscopy while using 20% as cut off.
Categorical score: manual assessment versus automated assessment: cutoff 14%

Optical microscopy versus quantitative analysis: cutoff 14%

<table>
<thead>
<tr>
<th></th>
<th>Observer A</th>
<th>Observer B</th>
<th>Observer C</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>88,2</td>
<td>86,8</td>
<td>90,8</td>
</tr>
<tr>
<td>Low</td>
<td>25</td>
<td>11,8</td>
<td>20,8</td>
</tr>
</tbody>
</table>

Computer screen versus quantitative analysis: cutoff 14%

<table>
<thead>
<tr>
<th></th>
<th>Observer A</th>
<th>Observer B</th>
<th>Observer C</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>85,5</td>
<td>96,1</td>
<td>78,9</td>
</tr>
<tr>
<td>Low</td>
<td>29,2</td>
<td>37,5</td>
<td>8,3</td>
</tr>
</tbody>
</table>

Agreement between manual and automated assessment of Ki 67 indexes using 14% as cutoff point: moderate to substantial: kappa values ranging from 0.585 to 0.680 on optical microscopy and 0.533 to 0.640 (slightly lower) on virtual slides (p value <0.01).

Overall tendency was to overestimate the Ki 67 index using both methods of manual ki 67 assessment in comparison to automated assessment of Ki 67 values.
Categorical score: manual assessment versus automated assessment: cutoff 20%

**Optical microscopy** versus quantitative analysis: cutoff 20%

<table>
<thead>
<tr>
<th>Observer</th>
<th>High QA</th>
<th>Low QA</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Observer A</td>
<td>80.9</td>
<td>15.6</td>
</tr>
<tr>
<td>Low Observer A</td>
<td>84.4</td>
<td>19.1</td>
</tr>
<tr>
<td>High Observer B</td>
<td>76.1</td>
<td>15.6</td>
</tr>
<tr>
<td>Low Observer B</td>
<td>84.4</td>
<td>23.9</td>
</tr>
<tr>
<td>High Observer C</td>
<td>80.9</td>
<td>12.5</td>
</tr>
<tr>
<td>Low Observer C</td>
<td>87.5</td>
<td>19.1</td>
</tr>
</tbody>
</table>

**Computer screen** versus quantitative analysis: cutoff 20%

<table>
<thead>
<tr>
<th>Observer</th>
<th>High QA</th>
<th>Low QA</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Observer A</td>
<td>73.5</td>
<td>15.6</td>
</tr>
<tr>
<td>Low Observer A</td>
<td>84.4</td>
<td>26.5</td>
</tr>
<tr>
<td>High Observer B</td>
<td>86.8</td>
<td>18.8</td>
</tr>
<tr>
<td>Low Observer B</td>
<td>81.2</td>
<td>13.2</td>
</tr>
<tr>
<td>High Observer C</td>
<td>63.2</td>
<td>6.2</td>
</tr>
<tr>
<td>Low Observer C</td>
<td>93.8</td>
<td>36.8</td>
</tr>
</tbody>
</table>

Agreement between manual and automated assessment of Ki 67 indexes using 20% as cutoff point: moderate to substantial: kappa values ranging from 0.558 to 0.636 on optical microscopy and 0.479 to 0.664 (slightly lower) on virtual slides (p value <0.01).

Overall tendency was to underestimate the Ki 67 index using both methods of manual ki 67 assessment in comparison to automated assessment of Ki 67 values.
Inter-modalities agreement:

2 observers examined the cases on optical microscopy then on virtual slides.

Inter-modalities agreement was almost perfect with CC estimated at 0.822 for observer B and 0.903 for observer C (p value <0.01).
Comparison between the digital quantitative analysis of Ki67 proliferative index for all the snap shots taken from all cases versus done for the WSIs
Conclusions

- Agreement between manual assessment of Ki 67 index and automated assessment was only substantial with slightly better agreement between optical microscopy results and automated quantitative analysis in comparison to virtual slides results.

- While applying the cut off points: 14 and 20%, agreement between manual and automated assessment of Ki 67 indexes using both cutoff points was slightly better with optical microscopy results in comparison to virtual slides.

- The inter-observers agreement using both manual methods as well as the agreement between the manual Ki 67 assessment results and the automated quantitative assessment of Ki 67 were better while using 14% as cutoff point than while using 20% as a cutoff point.
The 3 observers overestimated the values of Ki 67 while using 14% as a cut off points with both manual methods in comparison to quantitative analysis results but underestimated the Ki 67 values while 20% was used as a cut off points.

The results of ki 67 assessment for the group of patient with ki 67 values between 14-20%, the group in which most controversial cut offs are located for making clinical decisions, were highly discordant between the 3 observers and quantitative analysis.
However, without standardization of methodology, these cutoffs have limited value outside of the studies from which they were derived and the centers that performed them.

Many studies showed that automated quantitative analysis results were more correlated with clinico-pathological characteristics of breast tumors as well as prognostic factors and that Digital Image Analysis (DIA) methods of scoring Ki67 outperformed pathologist’s manual scores in terms of sensitivity and specificity for the Luminal B subtype classification while compared to PAM50 gene expression assays.*
The Comparison between the digital quantitative analysis of Ki 67 proliferative index for all the snap shots taken from all cases versus that obtained from automated quantitative analysis of the WSIs revealed that although the use of later method is much easier and saving time but results of the digital quantitative for the snap shots were more closer to the manual results and consequently much acceptable to the pathologists.

While that results obtained from the quantitative analysis of Ki 67 proliferative index for WSIs were not uniform [much higher in some cases and much lower in others] in comparison with that obtained from the quantitative analysis of the snapshots. Consequently it was much away from the manual results and not acceptable for the pathologists.
This evident difference between the two automated quantitative analysis of the cases [digital quantitative analysis of Ki 67 proliferative index for all the snap shots and the digital quantitative analysis of Ki 67 of the WSIs] can be explained by the bias occurred during the digital quantitative analysis of Ki 67 done for the WSIs due to the areas of tumoural necrosis & inflammatory cellular infiltrate which can add false positive results.

So we concluded that the use of digital quantitative analysis of Ki 67 done for the WSI needs using some computerized filters to exclude these misleading areas which may give false results.
Automated assessment of Ki 67 value would appear to be a potential easy-to-use tool for a robust and standardized fully automated Ki 67 scoring replacing the widely criticized current manual evaluation especially that a growing number of ready-to-use systems are offered on the market making the DIA an accessible, simple option with superior reproducibility.

Using Automated Quantitative Analysis of the WSI is easier & much faster but needs some modifications to be more accurate & acceptable
Faculty of medicine, Kasr El-Eini Hospital [1823-2016]
THANK YOU