Practical approaches to automated digital image analysis of Ki-67 labeling index in 997 breast carcinomas and causes of discordance with visual assessment

Ah-Young Kwon[1], **Ha Young Park**[2], **Jiyeon Hyeon**[3], **Seok Jin Nam**[4], **Seok Won Kim**[4], **Jeong Eon Lee**[4], **Jong-Han Yu**[4], **Se Kyung Lee**[4], **Soo Youn Cho**[5]*, **Eun Yoon Cho**[5]*

1 Department of Pathology, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam, Gyeonggi-do, Republic of Korea, 2 Department of Pathology, Busan Paik Hospital, Inje University College of Medicine, Busan, Republic of Korea, 3 Department of Pathology, Incheon St Mary's hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea, 4 Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea, 5 Department of Pathology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

☯ These authors contributed equally to this work.
* eunyoon.cho@samsung.com; sooyoun.cho@samsung.com

Abstract

The Ki-67 labeling index (LI) is an important prognostic factor in breast carcinoma. The Ki-67 LI is traditionally calculated via unaided microscopic estimation; however, inter-observer and intra-observer variability and low reproducibility are problems with this visual assessment (VA) method. For more accurate assessment and better reproducibility with Ki-67 LI, digital image analysis was introduced recently. We used both VA and automated digital image analysis (ADIA) (Ventana Virtuoso image management software) to estimate Ki-67 LI for 997 cases of breast carcinoma, and compared VA and ADIA results. VA and ADIA were highly correlated (intraclass correlation coefficient 0.982, and Spearman's correlation coefficient 0.966, p<0.05). We retrospectively analyzed cases with a greater than 5% difference between VA and ADIA results. The cause of these differences was: (1) tumor heterogeneity (98 cases, 56.0%), (2) VA interpretation error (32 cases, 18.3%), (3) misidentification of tumor cells (26 cases, 14.9%), (4) poor immunostaining or slide quality (16 cases, 9.1%), and (5) Estimation of non-tumor cells (3 cases, 1.7%). There were more discrepancies between VA and ADIA results in the group with a VA value of 10–20% compared to groups with <10% and ≥20%. Although ADIA is more accurate than VA, there are some limitations. Therefore, ADIA findings require confirmation by a pathologist.

Introduction

Breast cancer is one of the most frequent malignancies in women. Many studies have sought to improve treatment outcomes for breast cancer, and molecular studies play a critical role in prognosis. Intensive molecular studies have made it possible to classify breast carcinomas,

leading to improvements in treatment, prognosis prediction, and outcomes. Currently, the molecular classification of breast carcinomas can be easily confirmed according to estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor 2 (HER2), or Ki-67 labeling index (LI) status. Immunohistochemistry (IHC) is widely used to determine the expression of these markers. Among them, Ki-67 LI is a parameter for molecular classification and prognostic assessment[1–4]. For instance, in ER-positive and HER2-negative breast cancers in particular, the classification of subtypes is dependent on Ki-67 LI: tumors with low Ki-67 LI are classified into the luminal A group and those with high Ki-67 LI into the luminal B group[5]. Recurrence rate, prognosis, and therapeutic recommendations differ according to subtypes.

Traditionally, Ki-67 LI is estimated by visual observation. Despite its importance, Ki-67 LI has high inter-observer and/or intra-observer variability, and low reproducibility[6–8]. Several methods, such as a five-grade scale, have been suggested to resolve this problem[9–11]; however, low reproducibility remains an issue. Recently, computer-assisted image analysis has been used to achieve higher reproducibility of IHC results. Automated digital image analysis (ADIA) of Ki-67 LI in breast cancers obtains high quality data[12]. Several studies have observed ADIA to yield more reproducible and accurate measurement[13–16], and its application is being implemented in a clinical setting.

We performed visual assessment (VA) and ADIA simultaneously for breast carcinoma cases for approximately seven months. VA and ADIA were shown have their own strengths and weaknesses. In this study, we compared Ki-67 LI between by VA and by ADIA, and analyzed the causes of discrepancies. We sought to share the advantages and limitations of AIDA based on our experience.

Materials and methods

Patients and tissue specimens

This was a retrospective study conducted at a single institution. We collected all excised or biopsied specimens from patients diagnosed with breast carcinoma who underwent Ki-67 LI analysis at the Samsung Medical Center from December 2015 to June 2016. A total of 997 consecutive breast cancer specimens from 964 patients were obtained. All samples were formalin-fixed, paraffin-embedded and processed in a pathology laboratory according to standardized institutional protocols. Clinicopathological data were obtained by reviewing clinical charts for age, specimen type, stage, histological diagnosis, nuclear grade, IHC profiles for ER, PR and HER2, and the results of HER2 silver in situ hybridization (SISH). The study was approved by the institutional review board at Samsung Medical Center, Seoul, Korea (IRB No.2016-09-099), and informed consent was waived. All data were fully anonymized when we collected them.

Immunohistochemical staining for Ki-67

IHC was performed with a Ventana automated immunostainer (Ventana, Tucson, AZ, USA). Tissue sections 4 μm thick were cut, dried, deparaffinized, rehydrated, and heated following a standard protocol. Primary Ki-67 antibody (MIB-1, DAKO, Denmark) was used at 1:200 dilution with the DAB detection system (Ventana) protocol.

Assessment of Ki-67 by visual assessment

For VA, all Ki-67-immunostained slides were evaluated independently by two of four pathologists (AY Kwon, EY Cho, SY Cho, or HY Park) during routine reading. At the time of the

evaluation, the pathologist was blinded to the previous estimated value. In cases of discrepancy between the two pathologists, an adjusted value was reached through consensus. Evaluation of Ki-67 nuclear positivity was calculated by unaided microscopic estimating to determine the percentage of tumor cells that were Ki-67 immunostaining-positive. If an invasive component was present, Ki-67 LI was estimated only in the invasive tumor component. If only an intraductal component was present, or the invasive area was less than 1 mm, Ki-67 LI was evaluated in the intraductal component. Using a protocol from the International Ki-67 in Breast Cancer Working Group[17], Ki-67 LI was evaluated in at least three fields at high-power magnification as the percentage of Ki-67 nuclear positivity. In the somewhat large tumor, we identified the homogeneity of Ki-67 staining at low-power magnification. In the homogeneously stained tumor, three randomly selected areas were used for evaluation. In the heterogeneously stained tumor, we evaluated the area including relatively hot spots, where the highest ratio of positive cells is shown. VA results for Ki-67 LI were estimated in 10% increments: if less than 10% was detected, it was subdivided into $< 1\%$, $< 5\%$, or 5–10%. The VA value used in this analysis was the median of the measurement range.

Assessment of Ki-67 by automated digital image analysis

For ADIA, each Ki-67 stained slide of 997 cases was digitally scanned using a Ventana iScan (Ventana) with a 200x objective. Before scanning, pathologists selected a region of interest (ROI) about 25–50 mm^2 in size (about one to two low-power field), including the area for VA assessment, especially hot spots. After scanning, the procedure, also called semi-manual image analysis, was performed as follows. Experienced physician assistants (PAs) selected some areas randomly in the ROI at 200x magnification. As seen in Fig 1, the PA-selected area is the green box. In this box, black lines were used to exclude the stromal area by the PA (Fig 1). After setting this area, Ventana Virtuoso image management software (ver.5.6, Ventana) was used to
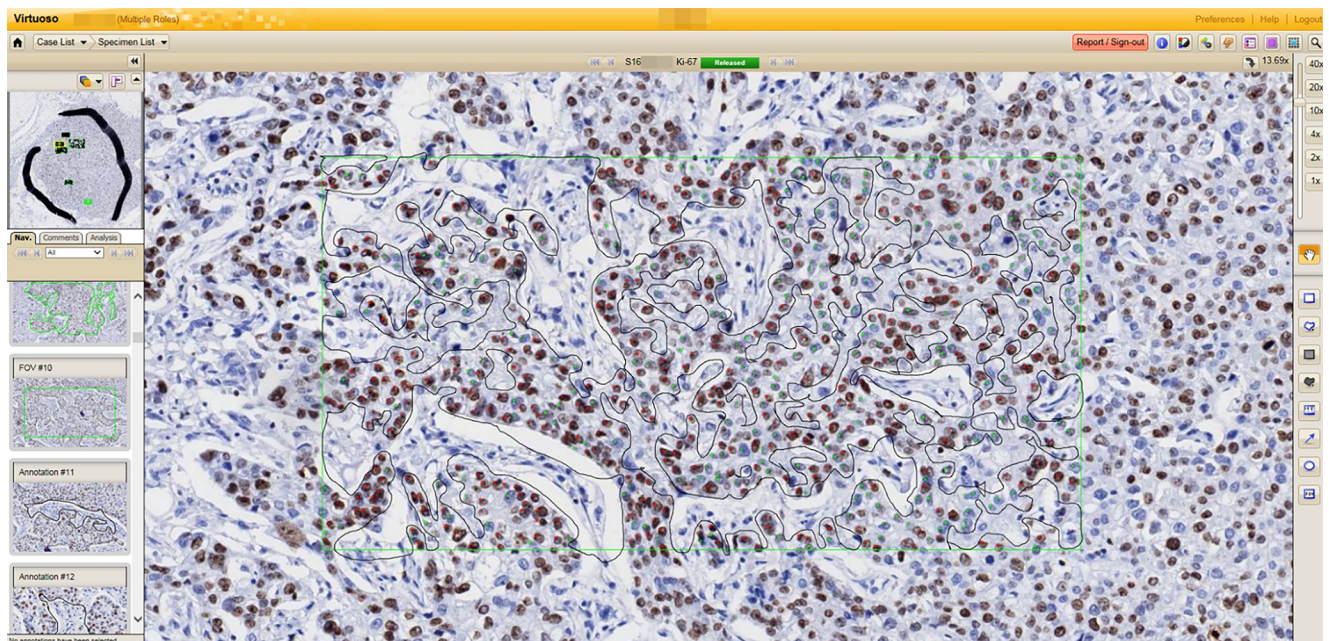
Fig 1. Automated digital image analysis by Ventana Virtuoso image analysis software. Each slide is scanned and several areas are selected. A green line surrounds the selected area, and a black line excluded area. The selected areas are analyzed automatically, marking stained tumor cells with red dots and non-stained tumor cells with green dots. The non-tumor cells are excluded in analysis automatically.

https://doi.org/10.1371/journal.pone.0212309.g001

evaluate positive and negative stained tumor cells automatically. Ki-67-stained tumor nuclei were labeled and marked with DAB dye automatically by the software in the selected areas as red dots, while negative tumor nuclei remained the blue-color of hematoxylin staining, and were denoted by green dots. By repeating the above procedure, at least 1,000 tumor cells were evaluated according to recommendations of the Ventana software. In the practice, we tried to evaluated at least 1,500 tumor cells, an average of 3,000 cells. If tumor cells accounted for less than 1,000, all cells were evaluated. In ADIA, like VA, the invasive component was also assessed. When there was no invasive component or less than 1 mm, ADIA was evaluated in the carcinoma in situ. All ADIA results were stored on a server provided by Ventana. ADIA was performed by a PA and confirmed by a pathologist.

Investigation of discrepancies between visual assessment and automated digital image analysis

We identified samples with a discrepancy between ADIA value and the median VA value greater than 5%, or an ADIA value out of the VA range. We sought to determine the cause of these discrepancies, and the limitations of VA or ADIA. By comparing Ki67-stained slides and the corresponding digital images, we surveyed and categorized the causes of these discrepancies.

Statistical analysis

For analysis of the association between VA and ADIA, Pearson's and Spearman's rank correlations and intraclass correlation coefficient (ICC) were used. Correlations between clinico-pathological parameters and differences in Ki-67 LI values were analyzed by chi-square test or Fisher's exact test. Statistical analysis was performed using SPSS 16.0 for Windows (SPSS Inc., Chicago, IL, USA).

Results

Patient and tumor characteristics

We analyzed 997 breast carcinomas from 964 patients, with 22 patients having undergone bilateral mastectomy for bilateral breast carcinomas, and 11 patients undergoing Ki-67 LI testing of both a biopsy specimen of the breast or lymph node and a mastectomy specimen. The characteristics of 964 patients and 997 tumors are summarized in Table 1, and all data is provided in S1 File.

Of the patients, 47.1% (454/964) were 50 years of age or older (23–85 years, median 49 years). Specimens were 14 core biopsies, 8 mammotome biopsies, 27 excisions, and 598 samples from partial mastectomy, and 350 from total mastectomy. Excluding 11 cases with repeat examination and 6 cases of exams where only metastatic lesions were examined, histological type was analyzed for a total of 980 breast carcinomas (Table 1). For invasive carcinomas, tumor size was 10 μm (microinvasive carcinoma with extensive intraductal component) to 14 cm (median, 1.7 cm). For 979 breast carcinomas of assessable nuclear grade, 13.1% were low nuclear grade, 59.9% were intermediate, and 27.1% were high. In addition, 78.8% of tumors were ER positive, 70.6% were PR positive, and 18.5% were HER2 positive. Samples from 60 patients (6.2%) were examined by IHC after neoadjuvant chemotherapy.

Correlation between visual assessment and automated digital image analysis for Ki-67 LI

The median VA value from an experienced pathologist was 22.86% (0.5%–95.0%, median 15.00%) and the median ADIA value was 23.43% (0.09%–94.71%, median 15.51%). VA and

Table 1. Clinicopathologic characteristics of 997 cases with breast carcinoma of 964 patients.

Clinicopathologic parameters		n (%)
Age (years)	< 50	510 (52.9%)
	≥ 50	454 (47.1%)
Specimen	Core biopsy	14 (1.4%)
	Mammotome biopsy	8 (0.8%)
	Excision	27 (2.7%)
	Lumpectomy (partial mastectomy)	598 (60.0%)
	Total mastectomy	350 (35.1%)
T stage (11 cases, repeat exam) (10 cases, not assessable)	Tis	158 (16.2%)
	T1	506 (51.8%)
	T2	267 (28.3%)
	T3	42 (4.3%)
	T4	3 (0.3%)
N stage (11 cases, repeat exam) (65 cases, not assessable)	Metastasis present (N1, 2, 3)	280 (30.4%)
	Metastasis absent (N0)	641 (69.6%)
Histologic diagnosis (11 cases, repeat exam) (6 cases, not assessable)	Invasive ductal carcinoma	695 (70.9%)
	Ductal carcinoma in situ	149 (15.2%)
	Invasive lobular carcinoma	57 (5.8%)
	Lobular carcinoma in situ	6 (0.6%)
	Mixed invasive ductal and lobular carcinoma	8 (0.8%)
	Mucinous carcinoma	30 (3.1%)
	Metaplastic carcinoma	12 (1.2%)
	Invasive micropapillary carcinoma	9 (0.9%)
	Tubular carcinoma	5 (0.5%)
	Encapsulated papillary carcinoma	4 (0.4%)
	Solid papillary carcinoma	3 (0.3%)
	Giant cell carcinoma	2 (0.2%)
Nuclear grade (11 cases, repeat exam) (7 cases, not assessable)	Low	128 (13.1%)
	Intermediate	587 (59.9%)
	High	265 (27.0%)
Immuno-histochemical stain (11 cases, repeat exam)	ER positive	777 (78.8%)
	ER negative	209 (21.2%)
	PR positive	696 (70.6%)
	PR negative	290 (29.4%)
	HER2 positive	183 (18.6%)
	HER2 equivocal	39 (4.0%)
	HER2 negative	764 (77.5%)

https://doi.org/10.1371/journal.pone.0212309.t001

ADIA evaluations were highly correlated (ICC 0.982, 95% confidence interval [CI] 0.980–0.984, Pearson's correlation coefficient 0.982 [$p<0.05$], and Spearman's correlation coefficient 0.966 [$p<0.05$]).

Comparison of the differences between visual assessment and automated digital image analysis by the groups

The VA cutoff value of Ki-67 LI for molecular classification and decision making is 14%. Therefore, we classified all cases according to the VA value of Ki-67 LI as <10%, 10%-20%,

Table 2. The intraclass correlation coefficient (ICC) and Spearman's correlation coefficient, stratified by visual assessment values.

Groups	n	ICC (95% CI)	p-value	Spearman's ρ	p-value
<10%	350	0.581 (0.507–0.646)	<0.001	0.698	<0.001
10%-20%	264	0.407 (0.301–0.503)	<0.001	0.546	<0.001
≥20%	383	0.965 (0.957–0.971)	<0.001	0.947	<0.001

https://doi.org/10.1371/journal.pone.0212309.t002

and ≥20%. The 10%-20% and <10% (ICC 0.407 and 0.581, respectively) groups showed poorer consistency than the of ≥20% group (ICC 0.947, Table 2)

Cause of discrepancies between visual assessment and automated digital image analysis

We retrospectively analyzed 175 cases with discrepancies greater than 5% between the median VA value of VA and ADIA results or with ADIA values beyond the VA range.

Causes of discrepancies were as follows: (1) tumor heterogeneity (98 cases, 56.0%); (2) VA interpretation error (32 cases, 18.3%); (3) misidentification of tumor cells (26 cases, 14.9%); (4) poor quality immunostaining or mounting, including weak staining, cytoplasmic staining, or problems with fixation or paraffin block-cutting (16 cases, 9.1%); and (5) non-tumor cells such as inflammatory cells and/or stromal cells included in estimation (3 cases, 1.7%). Analysis of causes of errors is summarized in Table 3.

Statistical analysis of discrepancies and clinicopathological parameters

We examined correlations according to type of specimen (biopsy, excision, mammotome, and mastectomy), histological subtype, tumor size, microinvasiveness, nuclear grade, neoadjuvant chemotherapy, PA who performed ADIA, or pathologist who performed VA. Biopsied specimens showed significantly larger discrepancies between VA and ADIA than mastectomy specimens (p = 0.035). Also, intermediate nuclear grade tumors had significantly larger discrepancies between VA and ADIA than tumors of high (p = 0.018) and low nuclear grade (p = 0.048). Other parameters were not correlated with discrepancies in Ki-67 LI.

Discussion

Ki-67 LI is an important marker for classification, treatment planning and prognosis of various malignant tumors. In breast carcinoma, as in other carcinomas, high Ki-67 LI is associated with worse outcomes[6, 18]. Although Ki-67 LI is important clinically, traditional evaluation depends on subjective assessment by pathologists. Relatively inaccurate and low-reproducible VA indexes[19] can affect the association between Ki-67 index and prognosis. ADIA is one solution for accurate and reproducible assessment of Ki-67 LI, that was recently introduced [16]. Using ADIA, relatively accurate estimation of stained tumor cells is possible in a few seconds. Permanent digital data is generated and repeat assessment is available at any time. ADIA

Table 3. Causes of Ki-67 LI discrepancies between visual assessment and automated digital image analysis.

Cause of discrepancy	n (%)
Tumor heterogeneity	98 (56.0%)
Visual assessment interpretation error	32 (18.3%)
Misidentification of tumor cells	26 (14.9%)
Poor immunostaining or slide quality	16 (9.1%)
Non-tumor cells estimated	3 (1.7%)

https://doi.org/10.1371/journal.pone.0212309.t003

offers little fear of loss of staining, labeling or whole slides. We used ADIA for breast cancer evaluation for about seven months. Through that experience, we recognized the practical limitations of ADIA use.

Previous studies focused mainly on correlation between VA and ADIA results, including area-setting methods[20, 21], ROI size and the number of cells[22], dual staining with cytokeratin and Ki-67[14], and other applications[23–25]. However, in this study, we investigated the limitations of ADIA in practice as well as comparing VA and ADIA, and we recommended methods to overcome the potential problems. We analyzed 175 cases with some discrepancies between VA and ADIA results.

Correlation analysis showed high concordance between VA and ADIA values. This result may indicate consistency; however, considering the effort to develop and expand ADIA use, it may be disappointing. Previous published studies showed good concordance between VA and ADIA, not only in breast cancer[26, 27], but also in other cancers[28–30]. Therefore, ADIA's main advantages are reproducibility[15, 16], permanence, and data storage. Also, ADIA is useful when it required specific values; for example, "14%" of breast cancers.

Through our experience and analysis of 175 cases, we identified the following possible problems. First, heterogeneity is the most frequent cause of discordance between VA and ADIA. An example of heterogeneity is presented in Fig 2A. Whereas pathologists intuitively judged the wider area during VA, the ADIA area was smaller.

Heterogeneity of Ki-67 labeling is an ongoing issue. In a study by Koopman et al.[31], heterogeneity was also the most common reason for discrepancies between manual counting and ADIA. To overcome heterogeneity, it is commonly recommended that Ki-67 LI be assessed in peripheral tumor areas with higher LI; however, area selection remains the most important aspect of assessment of Ki-67 LI. Some studies have worked to address this problem, by investigating selection method, ROI size, and the number of cells. Ki-67 LI is dependent on ROI size, and the number of cells estimated, which is a critically important parameter in Ki-67 quantification by ADIA. Christgen et al. showed that median Ki-67 LI decreased from 55 to 15% depending on ROI size and suggested that a minimal ROI of 500 cells is an acceptable compromise[22]. Our findings showing that biopsied specimen analysis yielded greater discrepancies between VA and ADIA results than analysis of mastectomy specimens support Christgen et al.'s results. Because biopsied specimens usually have relatively few tumor cells that are interrupted by stroma, smaller numbers of tumor cells are counted and discrepancies between VA and ADIA may be more prevalent. Analyzing the entire slide may be a way to resolve the heterogeneity-problem. Laurinavicius et al. assessed Ki-67 expression on whole-slide images by digital image analysis and demonstrated that Ki-67 is an independent prognostic indicator of overall survival[32]. Zhong et al. investigated an average score method, assessing at least three selected areas on an entire digital slide, with comparison to a hot-spot score method focusing on hot spot areas[33]. This average score method led to slightly better agreement between VA and ADIA results than the hot-spot score method, but differences were extremely small. Whether values from whole-slide or hot-spot analysis by ADIA are more useful for prognosis remains unclear. Further studies analyzing ADIA results and patient prognosis are needed to improve the precision of scoring and to develop new cutoff values for Ki-67 LI for breast cancer subtyping. The area for ADIA was selected in which was previously assessed by VA, and estimated at least 1,500 and an average of 3,000 tumor cells by ADIA; however, heterogeneity remains a problem in ADIA. Area selection is always accompanied by selection bias unless a whole slide section is measured. Therefore, in practice, we recommend that the basic protocol for selecting areas is followed, and that a large number of tumor cells, at least 1,000 cells, are counted. For cases with extremely heterogeneous staining, pathologists can confirm the
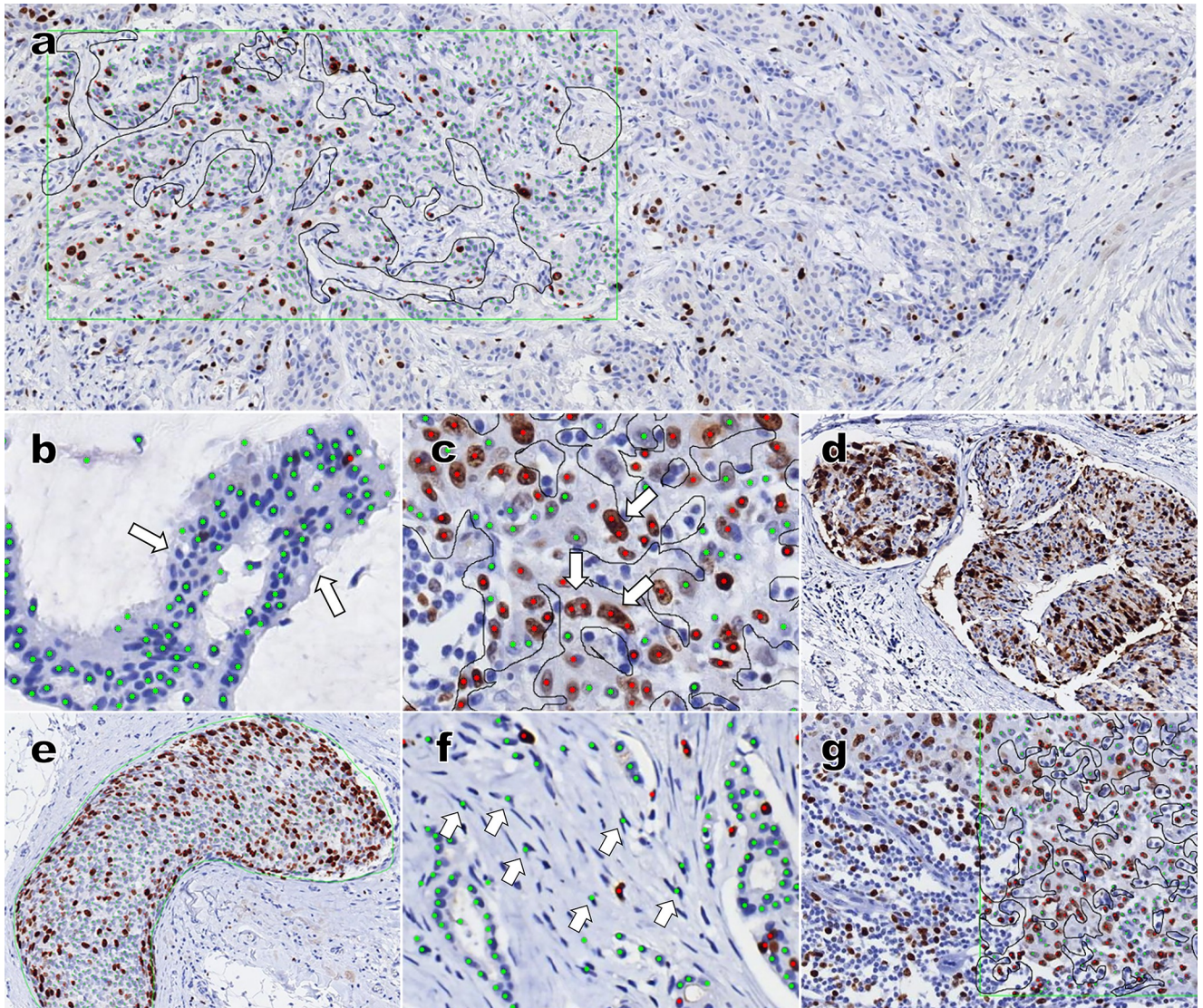
Fig 2. The causes of discrepancy between two visual assessment and automated data image analysis. (a) Heterogeneity of Ki-67 immunostaining. The left side green-lined box had a higher Ki-67 labeling index (LI) than the right side, indicating a difference between visual assessment and automated digital image analysis. (b—c) Misidentification of tumor cells. (b) Some tumor cells (arrow) were not recognized due to limitation of the automated algorithm. (c) Pleomorphic large cells or dumbbell-shaped tumor cells were estimated as two or more cells. (d—e) Poor Ki-67 immunostaining quality. (d) Poor tissue fixation led to poor Ki-67 staining quality and interfered with automatic image analysis. (e) Changing the tissue block and re-staining for Ki-67 led to clear staining and image analysis. (f—g) Erroneous recognition of non-tumor cells as Ki-67-negative tumor cells. (f) Some stromal cells were assessed as Ki-67-negative tumor cells. (g) Areas with stromal cells and/or inflammatory cells should be excluded; tumor cells were selected (black line) to exclude non-tumor cells as much as possible.

findings using VA. Pathologists need to carefully consider ROI selection, and may try to estimate more areas multiple times with ADIA.

The second most common cause of discrepancy was VA interpretation error. We were aware of interpretation errors at the time of glass slide review. The main reason for errors was intuitive or partial measurements, or mistakes in recording. In these 32 cases, we were able to use ADIA to correct human error.

The third cause of differences was misidentification of tumor cells. In some cases, Ki-67-negative tumor cells were not marked (Fig 2B). This might have been due to weak

hematoxylin staining and/or a limitation of the automated software algorithm. Another possibility in a few cases was a pleomorphic, large cell or a dumbbell-shaped tumor cell being estimated as two or more cells (Fig 2C). When we recognized these errors during confirmation, we requested that PAs repeat ADIA. Therefore, pathologists must be verified for Ventana scanning images and judging if ADIA analysis is appropriate. If needed, re-analysis of other areas must be performed. We think that manual marking or deleting dots from images may be helpful; however, this is currently not available in Ventana software, which must follow the algorithm provided.

The fourth cause of discrepancies was poor Ki-67 immunostaining quality. Problems with tissue fixation, tissue processing, and/or thick slices because of abundant fat tissue, were included in this category. The spindling morphology of tumor cells due to poor preparation or fixation could not be used to differentiate tumor cells from stromal cells on ADIA (Fig 2D). A few samples showed clearing of nuclei or cytoplasmic Ki-67-staining, which interfered with automatic image analysis. Through experience we realized that these slides made ADIA analysis difficult. We began to evaluate slide quality during VA, and re-stained for Ki-67, recut slides, or changed tissue blocks as necessary (Fig 2E). There is still the potential for error due to these issues, and pathologists and PAs should be fully aware of these concerns.

Another cause of discrepancy was misidentification of non-tumor cells as Ki-67-negative tumor cells. ADIA can recognize stromal cells and/or inflammatory cells as tumor cells, so they are assessed and counted as Ki-67-negative tumor cells (Fig 2F). Therefore, it was helpful that non-tumor cells were excluded in the estimated area as much as possible, by using the black line in Fig 2G. However, non-tumor cells, such as lymphocytes infiltrated in carcinomas, were unavoidably included in some estimates. Harvey et al. and Roge et al.[14, 34] suggested double-staining of cytokeratin and Ki-67 to exclude stromal and/or inflammatory cells. The estimation protocol we used, Ventana Virtuoso image management software, did not involve double-staining. We expect that double-staining could be helpful, but only 1.7% of our cases involved this error. In other words, while this problem has the potential to affect Ki-67 LI measurement, it was not found to be a major concern in our protocol. A new measurement method using double-staining could overcome this issue, enabling more accurate measurements.

In the situation where the discrepancy between VA and ADIA was more than 5%, ADIA tended to measure relatively higher than VA; 116 cases showed higher Ki-67 LI in ADIA and 59 cases in VA. We analyzed the relative values of VA and ADIA for each cause. In cases of heterogeneity, 69 cases (70.4%) showed higher ADIA than VA, indicating a tendency toward relatively high Ki-67 LI on ADIA when there was heterogeneity. The smaller and more compact the ROI, the more Ki-67 LI was expected to increase. In cases of VA error, VA was higher in half cases (16/32, 50.0%) and half cases showed higher ADIA. In cases of misidentification of tumor cells, there were 22 cases (84.6%) with higher ADIA values and 4 cases (15.4%) with higher VA values. It is believed that tumor cells without Ki-67 labeling were often not counted, or stained tumor cells were counted multiple times. In the cases of poor slide quality, half (8/16, 50.0%) showed higher VA values and the other half showed higher ADIA values. The three cases of non-tumor cell estimation consisted of 2 cases with higher VA and 1 case with higher ADIA.

Because 14% is widely used as a cutoff for molecular classification and a prognostic factor [35], we employed three classification groups, <10%, 10%-20%, and ≥20%, stratified by VA values. We compared the concordance between VA and ADIA among the three groups. As with previous studies[33], the intermediate Ki-67 LI group (10%-20%) showed relatively weak concordance between VA and ADIA. In addition, we analyzed ADIA values of samples with 10%-20% of VA Ki-67 LI, which included 198 cases (19.9%). Additionally, we selected the

cases that were ER/PR positive and HER2 negative, because Ki67 LI is an important factor in luminal A and B subtyping. A total of 560 cases were culled. Of these, in 4 cases, the VA value was 5%-15%, but ADIA values were over 14% higher; in other words, according to the VA values, the cancers were luminal A subtype, but according to ADIA values, they were luminal B subtype. When the VA value was 10%-20%, there were more luminal B subtypes (Ki-67 LI$\geq$14%, 91/131 cases) than luminal A subtypes (Ki-67 LI<14%, 40/131 cases). Thus, when the VA value is close to 14%, there is weak concordance, and it is not easy to determine the subtype. In these cases, ADIA will be certainly helpful. Even when using ADIA, there may still be cases for which classification is difficult. We recommend repeating measurement using more cell counts or a bigger area, and further studies on prognosis and subtype classification according to new measurement values are needed.

Carcinomas with intermediate nuclear grade showed larger differences in Ki-67 LI than carcinomas with low and high nuclear grades. We propose that high-grade nuclei were easily recognized due to their bigger size on both VA and ADIA. Also, high-grade tumors usually showed dense nesting patterns of tumor cells without accompanying stroma, with low overall heterogeneity because Ki-67 was diffusely stained. In low-grade tumors, Ki-67 LI was usually estimated to be less than 5% or 10%. Therefore, the discrepancy between VA and ADIA was not large among low-grade tumors.

VA has well-known limitations, especially inter-observer and intra-observer variability. Ki-67 LI by ADIA is expected to yield more reproducible and accurate values, and ADIA was a useful method in our experience for seven months of experience. Some limitations of ADIA were established from that experience, and as well as retrospective analysis of discrepancies between VA and ADIA values. In this paper, we analyzed the results of ADIA, and possible problems associated with the use of ADIA or VA. We sought to overcome these limitations, and have succeeded to some extent. We realized that final confirmation of ADIA data by experts was essential. These recommendations should lead to more accurate Ki-67 LI measurement, and these data will support prognostic analysis, which in turn will be helpful in tumor classification, determining prognosis, and treatment planning.

Supporting information

S1 File. Clinicopathologic characteristics of 997 cases with breast carcinoma of 964 patients, with VA and ADIA values.
(XLSX)

Author Contributions

Conceptualization: Eun Yoon Cho.

Data curation: Ah-Young Kwon.

Formal analysis: Ah-Young Kwon.

Investigation: Ah-Young Kwon, Ha Young Park, Soo Youn Cho, Eun Yoon Cho.

Resources: Seok Jin Nam, Seok Won Kim, Jeong Eon Lee, Jong-Han Yu, Se Kyung Lee.

Validation: Jiyeon Hyeon, Soo Youn Cho.

Writing – original draft: Ah-Young Kwon.

Writing – review & editing: Soo Youn Cho, Eun Yoon Cho.

References

1. Bouzubar N, Walker KJ, Griffiths K, Ellis IO, Elston CW, Robertson JF, et al. Ki67 immunostaining in primary breast cancer: pathological and clinical associations. British journal of cancer. 1989; 59(6):943–7. Epub 1989/06/01. PMID: 2472168; PubMed Central PMCID: PMCPMC2246720.

2. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. Annals of oncology: official journal of the European Society for Medical Oncology. 2009; 20 (8):1319–29. Epub 2009/06/19. https://doi.org/10.1093/annonc/mdp322 PMID: 19535820; PubMed Central PMCID: PMCPMC2720818.

3. Stuart-Harris R, Caldas C, Pinder SE, Pharoah P. Proliferation markers and survival in early breast cancer: a systematic review and meta-analysis of 85 studies in 32,825 patients. Breast (Edinburgh, Scotland). 2008; 17(4):323–34. Epub 2008/05/06. https://doi.org/10.1016/j.breast.2008.02.002 PMID: 18455396.

4. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. The Lancet Oncology. 2010; 11(2):174–83. Epub 2010/02/16. https://doi.org/10.1016/S1470-2045(09)70262-1 PMID: 20152769.

5. Gnant M, Harbeck N, Thomssen C. St. Gallen 2011: Summary of the Consensus Discussion. Breast care (Basel, Switzerland). 2011; 6(2):136–41. Epub 2011/06/03. https://doi.org/10.1159/000328054 PMID: 21633630; PubMed Central PMCID: PMCPMC3100376.

6. de Azambuja E, Cardoso F, de Castro G Jr, Colozza M, Mano MS, Durbecq V, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. British journal of cancer. 2007; 96(10):1504–13. Epub 2007/04/25. https://doi.org/10.1038/sj.bjc.6603756 PMID: 17453008; PubMed Central PMCID: PMCPMC2359936.

7. 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(5):e0125131. Epub 2015/05/02. https://doi.org/10.1371/journal.pone.0125131 PMID: 25932921; PubMed Central PMCID: PMCPMC4416820.

8. Fulawka L, Halon A. Ki-67 evaluation in breast cancer: The daily diagnostic practice. Indian Journal of Pathology and Microbiology. 2017; 60(2):177–84. https://doi.org/10.4103/IJPM.IJPM_732_15 PMID: 28631631

9. Hida AI, Bando K, Sugita A, Maeda T, Ueda N, Matsukage S, et al. Visual assessment of Ki67 using a 5-grade scale (Eye-5) is easy and practical to classify breast cancer subtypes with high reproducibility. Journal of clinical pathology. 2015; 68(5):356–61. Epub 2015/02/13. https://doi.org/10.1136/jclinpath-2014-202695 PMID: 25673730.

10. Mikami Y, Ueno T, Yoshimura K, Tsuda H, Kurosumi M, Masuda S, et al. Interobserver concordance of Ki67 labeling index in breast cancer: Japan Breast Cancer Research Group Ki67 ring study. Cancer science. 2013; 104(11):1539–43. Epub 2013/08/03. https://doi.org/10.1111/cas.12245 PMID: 23905924.

11. Romero Q, Bendahl PO, Ferno M, Grabau D, Borgquist S. A novel model for Ki67 assessment in breast cancer. Diagnostic pathology. 2014; 9:118. Epub 2014/06/18. https://doi.org/10.1186/1746-1596-9-118 PMID: 24934660; PubMed Central PMCID: PMCPMC4069280.

12. Abubakar M, Howat WJ, Daley F, Zabaglo L, McDuffus LA, Blows F, et al. High-throughput automated scoring of Ki67 in breast cancer tissue microarrays from the Breast Cancer Association Consortium. The journal of pathology Clinical research. 2016; 2(3):138–53. Epub 2016/08/09. https://doi.org/10.1002/cjp2.42 PMID: 27499923; PubMed Central PMCID: PMCPMC4958735.

13. Demichelis F, Barbareschi M, Dalla Palma P, Forti S. The virtual case: a new method to completely digitize cytological and histological slides. Virchows Archiv: an international journal of pathology. 2002; 441 (2):159–64. Epub 2002/08/22. https://doi.org/10.1007/s00428-001-0561-1 PMID: 12189506.

14. Harvey J, Thomas C, Wood B, Hardie M, Dessauvagie B, Combrinck M, et al. Practical issues concerning the implementation of Ki-67 proliferative index measurement in breast cancer reporting. Pathology. 2015; 47(1):13–20. Epub 2014/12/05. https://doi.org/10.1097/PAT.0000000000000192 PMID: 25474507.

15. Gudlaugsson E, Skaland I, Janssen EA, Smaaland R, Shao Z, Malpica A, et al. Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer. Histopathology. 2012; 61(6):1134–44. Epub 2012/09/12. https://doi.org/10.1111/j.1365-2559.2012.04329.x PMID: 22963617.

16. Voros A, Csorgo E, Kovari B, Lazar P, Kelemen G, Cserni G. The use of digital images improves reproducibility of the ki-67 labeling index as a proliferation marker in breast cancer. Pathology oncology research: POR. 2014; 20(2):391–7. Epub 2013/11/10. https://doi.org/10.1007/s12253-013-9708-z PMID: 24202856.

17. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. J Natl

Cancer Inst. 2011; 103(22):1656–64. Epub 2011/10/01. https://doi.org/10.1093/jnci/djr393 PMID: 21960707; PubMed Central PMCID: PMCPMC3216967.

18. Domagala W, Markiewski M, Harezga B, Dukowicz A, Osborn M. Prognostic significance of tumor cell proliferation rate as determined by the MIB-1 antibody in breast carcinoma: its relationship with vimentin and p53 protein. Clinical cancer research: an official journal of the American Association for Cancer Research. 1996; 2(1):147–54. Epub 1996/01/01. PMID: 9816101.

19. Chung YR, Jang MH, Park SY, Gong G, Jung WH, Korean Breast Pathology Ki-67 Study G. Interobserver Variability of Ki-67 Measurement in Breast Cancer. J Pathol Transl Med. 2016; 50(2):129–37. Epub 2016/02/16. https://doi.org/10.4132/jptm.2015.12.24 PMID: 26875758; PubMed Central PMCID: PMCPMC4804149.

20. Laurinavicius A, Plancoulaine B, Laurinaviciene A, Herlin P, Meskauskas R, Baltrusaityte I, et al. A methodology to ensure and improve accuracy of Ki67 labelling index estimation by automated digital image analysis in breast cancer tissue. Breast cancer research: BCR. 2014; 16(2):R35. Epub 2014/04/09. https://doi.org/10.1186/bcr3639 PMID: 24708745; PubMed Central PMCID: PMCPMC4053156.

21. Plancoulaine B, Laurinaviciene A, Herlin P, Besusparis J, Meskauskas R, Baltrusaityte I, et al. A methodology for comprehensive breast cancer Ki67 labeling index with intra-tumor heterogeneity appraisal based on hexagonal tiling of digital image analysis data. Virchows Archiv: an international journal of pathology. 2015. Epub 2015/10/21. https://doi.org/10.1007/s00428-015-1865-x PMID: 26481244.

22. Christgen M, von Ahsen S, Christgen H, Langer F, Kreipe H. The region-of-interest size impacts on Ki67 quantification by computer-assisted image analysis in breast cancer. Hum Pathol. 2015; 46 (9):1341–9. Epub 2015/07/25. https://doi.org/10.1016/j.humpath.2015.05.016 PMID: 26206765.

23. Bigras G, Dong WF, Canil S, Hugh J, Berendt R, Wood G, et al. New Robust and Reproducible Stereological IHC Ki67 Breast Cancer Proliferative Assessment to Replace Traditional Biased Labeling Index. Applied immunohistochemistry & molecular morphology: AIMM. 2017; 25(10):687–95. Epub 2016/04/20. https://doi.org/10.1097/pai.0000000000000371 PMID: 27093453; PubMed Central PMCID: PMCPMC5690316.

24. Fulawka L, Halon A. Proliferation Index Evaluation in Breast Cancer Using ImageJ and ImmunoRatio Applications. Anticancer research. 2016; 36(8):3965–72. Epub 2016/07/29. PMID: 27466501.

25. Maeda I, Abe K, Koizumi H, Nakajima C, Tajima S, Aoki H, et al. Comparison between Ki67 labeling index determined using image analysis software with virtual slide system and that determined visually in breast cancer. Breast Cancer. 2016; 23(5):745–51. Epub 2015/08/15. https://doi.org/10.1007/s12282-015-0634-7 PMID: 26271611.

26. Fasanella S, Leonardi E, Cantaloni C, Eccher C, Bazzanella I, Aldovini D, et al. Proliferative activity in human breast cancer: Ki-67 automated evaluation and the influence of different Ki-67 equivalent antibodies. Diagnostic pathology. 2011;6 Suppl 1:S7. Epub 2011/04/22. https://doi.org/10.1186/1746-1596-6-S1-S7 PMID: 21489202; PubMed Central PMCID: PMCPMC3073225.

27. Mohammed ZM, McMillan DC, Elsberger B, Going JJ, Orange C, Mallon E, et al. Comparison of visual and automated assessment of Ki-67 proliferative activity and their impact on outcome in primary operable invasive ductal breast cancer. British journal of cancer. 2012; 106(2):383–8. Epub 2012/01/19. https://doi.org/10.1038/bjc.2011.569 PMID: 22251968; PubMed Central PMCID: PMCPMC3261670.

28. Desmeules P, Hovington H, Nguile-Makao M, Leger C, Caron A, Lacombe L, et al. Comparison of digital image analysis and visual scoring of KI-67 in prostate cancer prognosis after prostatectomy. Diagnostic pathology. 2015; 10:67. Epub 2015/06/14. https://doi.org/10.1186/s13000-015-0294-0 PMID: 26070608; PubMed Central PMCID: PMCPMC4465166.

29. Reid MD, Bagci P, Ohike N, Saka B, Erbarut Seven I, Dursun N, et al. Calculation of the Ki67 index in pancreatic neuroendocrine tumors: a comparative analysis of four counting methodologies. Mod Pathol. 2015; 28(5):686–94. Epub 2014/11/22. https://doi.org/10.1038/modpathol.2014.156 PMID: 25412850; PubMed Central PMCID: PMCPMC4460192.

30. Tang LH, Gonen M, Hedvat C, Modlin IM, Klimstra DS. Objective quantification of the Ki67 proliferative index in neuroendocrine tumors of the gastroenteropancreatic system: a comparison of digital image analysis with manual methods. The American journal of surgical pathology. 2012; 36(12):1761–70. Epub 2012/10/03. https://doi.org/10.1097/PAS.0b013e318263207c PMID: 23026928.

31. Koopman T, Buikema HJ, Hollema H, de Bock GH, van der Vegt B. Digital image analysis of Ki67 proliferation index in breast cancer using virtual dual staining on whole tissue sections: clinical validation and inter-platform agreement. Breast Cancer Res Treat. 2018; 169(1):33–42. Epub 2018/01/20. https://doi.org/10.1007/s10549-018-4669-2 PMID: 29349710; PubMed Central PMCID: PMCPMC5882622.

32. Laurinavicius A, Plancoulaine B, Rasmusson A, Besusparis J, Augulis R, Meskauskas R, et al. Bimodality of intratumor Ki67 expression is an independent prognostic factor of overall survival in patients with invasive breast carcinoma. Virchows Archiv: an international journal of pathology. 2016; 468(4):493–502. Epub 2016/01/29. https://doi.org/10.1007/s00428-016-1907-z PMID: 26818835.

33. Zhong F, Bi R, Yu B, Yang F, Yang W, Shui R. A Comparison of Visual Assessment and Automated Digital Image Analysis of Ki67 Labeling Index in Breast Cancer. PloS one. 2016; 11(2):e0150505. Epub 2016/03/02. https://doi.org/10.1371/journal.pone.0150505 PMID: 26928407; PubMed Central PMCID: PMCPMC4771161.

34. Roge R, Riber-Hansen R, Nielsen S, Vyberg M. Proliferation assessment in breast carcinomas using digital image analysis based on virtual Ki67/cytokeratin double staining. Breast Cancer Res Treat. 2016; 158(1):11–9. Epub 2016/06/11. https://doi.org/10.1007/s10549-016-3852-6 PMID: 27283833.

35. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and progno-sis of patients with luminal B breast cancer. J Natl Cancer Inst. 2009; 101(10):736–50. Epub 2009/05/14. https://doi.org/10.1093/jnci/djp082 PMID: 19436038; PubMed Central PMCID: PMCPMC2684553.