



Image Analysis of HER2 Immunohistochemical Staining of Surgical Breast Cancer Specimens

Kwangil Yim¹, Hong Sik Park², Dong Min Kim², Youn Soo Lee², and Ahwon Lee^{2,3}

¹Department of Hospital Pathology, Uijeongbu St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Uijeongbu;

²Department of Hospital Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul;

³Cancer Research Institute, The Catholic University of Korea, Seoul, Korea.

Purpose: Trastuzumab is an effective treatment for human epidermal growth factor receptor 2 (HER2)-amplified breast cancers. We sought to develop a simple protocol for HER2 image analysis of breast cancer specimens.

Materials and Methods: In a preliminary test, we found that at least 1000 tumor cells need to be examined in the most strongly stained areas. Next, we evaluated the clinical usefulness of this established protocol of image analysis in 555 breast cancer patients. Results of the HER2 immunohistochemical (IHC) staining were compared between manual scoring and image analysis.

Results: The HER2 IHC results obtained by the image analysis method correlated well with those obtained by the manual scoring method (Cohen's kappa=0.830). Using the HER2 silver *in situ* hybridization (SISH) results as a gold standard, sensitivity values were 72.1% for manual scoring and 74.0% for image analysis; specificity values were 96.2% for manual scoring and 94.7% for image analysis; and accuracy values were 91.7% for manual scoring and 90.8% for image analysis. McNemar's test was applied to the results, and there were no statistically significant differences in sensitivity and specificity between the positive ($p=0.688$) and negative ($p=0.118$) SISH groups.

Conclusion: HER2 image analysis results were similar to those obtained via the manual scoring method, indicating that the use of image analysis can reduce assessment time and effort. We suggest that image analysis-based evaluation of 1000 tumor cells in the most strongly IHC-stained area, regardless of stroma content, is sufficient for determining HER2 expression levels in breast cancer specimens.

Key Words: EGFR, immunohistochemistry, SISH, breast cancer

INTRODUCTION

Human epidermal growth factor receptor 2 (HER2) is amplified in approximately 15–20% of breast carcinomas.^{1,2} Trastuzumab (Herceptin®; Genentech, South San Francisco, CA, USA) is a monoclonal antibody that binds to the extracellular

domain of the HER2 receptor, and patients with HER2 overexpression respond well to trastuzumab treatment.^{3,4}

HER2 tests are most commonly performed using immunohistochemistry (IHC), fluorescence *in situ* hybridization, or silver *in situ* hybridization (SISH) assays. Among them, IHC-based assessment is most widely used because it is inexpensive, is easy to undertake, and can be performed using familiar optical microscopy.⁵ With the integration of image analysis systems into clinical laboratories, objective evaluation, interpretation accuracy, and reproducibility can be improved.⁶ However, prior to using an image analysis system as part of routine testing, there are several important steps to consider, including ensuring consistent staining quality of IHC slides, providing a detailed protocol for selecting the analysis area, and developing a proven analysis algorithm. In this study, we used Algorithm HER2 (4B5) for image analysis along with VIRTUOSO software (Ventana Medical Systems, Tucson, AZ, USA) and an

Received: June 27, 2018 **Revised:** October 25, 2018

Accepted: November 30, 2018

Corresponding author: Ahwon Lee, MD, PhD, Department of Hospital Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Korea.
Tel: 82-2-2258-1621, Fax: 82-2-2258-1627, E-mail: klee@catholic.ac.kr

•The authors have no potential conflicts of interest to disclose.

© Copyright: Yonsei University College of Medicine 2019

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

iScan Coreo slide scanner (Ventana Medical Systems), an approach that was approved by FDA in 2011.

Most previous studies have been performed using biopsy specimens or tumor microarray. In those studies, all tumor cells were totally selected and analyzed. However, selecting all tumor cells included is time consuming and burdensome in routine practice for excised specimens. Therefore, simple and faster methods are needed for clinical settings. The aim of this study was to develop a simple protocol for HER2 image analysis that is not inferior to the HER2 manual scoring method for examining breast cancer specimens. We compared the results of the HER2 image analysis method to the results of the HER2 manual scoring method and to HER2 SISH results, which is considered the gold standard for assessing surgical specimens of breast cancer.

MATERIALS AND METHODS

Patients and study samples

Initially, in order to establish an image analysis-based method for HER2 measurement, we reviewed medical records for patients with confirmed invasive breast cancer who underwent surgical resections between January 2013 and December 2013 at Seoul St. Mary's Hospital in Seoul, Korea, and selected breast cancer patients with equivocal HER2 IHC-staining results but confirmed HER2 SISH results. Thirty-two patients from 376 assessed patients (8.51%) were included. In the next step, we evaluated the clinical usefulness of the newly established image analysis method in breast cancer patients who underwent surgical treatment at the same hospital between January 2011 and December 2012. During this period, 565 patients were histologically diagnosed with invasive breast cancer with both HER2 IHC and HER2 SISH tests to confirm HER2 status. Of these, 10 cases were excluded due to very small invasive foci or the absence of a slide to review. All specimens were routinely processed and diagnosed according to national and international guidelines. The present study was approved by the hospital's Institutional Review Board (approval KC17SESI0151).

Immunohistochemistry

The HER2 IHC-staining and HER2 SISH methods were performed following given instructions in the pathology laboratory manual. Briefly, for IHC analysis of HER2 expression, we selected a formalin-fixed, paraffin-embedded (FFPE) block containing both tumor and normal breast tissue to serve as an internal negative control. A HER2 SISH-positive breast cancer specimen was used as a positive control. The HER2 IHC staining was performed on an automated Ventana Benchmark XT platform using FDA-approved Ventana PATHWAY rabbit monoclonal antibody 4B5 clone and the Ventana ultraVIEW DAB Detection Kit (Ventana Medical Systems). HER2 expression was scored into one of four groups (0, 1+, 2+, 3+) accord-

ing to the 2013 American Society of Clinical Oncology/College of American Pathologists Guideline Update¹: 0, no staining observed or membrane staining that is incomplete, faint, or barely perceptible in $\leq 10\%$ of tumor cells; 1+, incomplete membrane staining that is faint or barely perceptible within $>10\%$ of tumor cells; 2+, circumferential membrane staining that is incomplete and/or weak/moderate and within $>10\%$ of tumor cells or with complete and circumferential membrane staining that is intense and within $\leq 10\%$ of tumor cells; and 3+, circumferential membrane staining that is complete and intense within $>10\%$ of tumor cells.

HER2 SISH analysis

HER2 SISH analysis was carried out on the same FFPE block that had been used for HER2 IHC. The SISH assay was performed with INFORM HER2 DNA probes (Ventana Medical Systems) according to the manufacturer's protocols. The probes were labeled with dinitrophenol and formulated for use with the Ventana ultraView SISH detection kit and the Ventana BenchMarkXT automated slide stainer. At least 20 tumor cells with positive black dot signals were counted in a homogeneous and contiguous population and then classified into one of three categories: negative if the HER2 copy number was <4.0 signals/cell, equivocal if there were 4–6 signals/cell, and positive if there were >6.0 signals/cell. When the first counting result was equivocal, signals of additional 20 cells were counted to determine the sample's HER2 status.

IHC scoring of HER2 using the image analysis system

We performed a preliminary analysis on a series of HER2 IHC-staining equivocal (2+) breast cancer specimens ($n=32$) to determine an optimal interpretation method. The 32 slides were scanned using an iScan Coreo slide scanner with a $20\times$ objective. All invasive cancer components of the entire slide image were carefully inspected to select foci of intense, thick, and more complete membrane staining [i.e., foci of view (FOVs)], and several ellipses with a constant area ($40457.64 \mu\text{m}^2$) were drawn until certain tumor cells were included (usually 6–10 ellipses) (Fig. 1). We analyzed the HER2 status of at least 500 tumor cells, 1000 tumor cells, and 2000 tumor cells in elliptically selected FOVs using Companion Algorithm HER2 (4B5) image analysis software. The results from the examined 500, 1000, and 2000 tumor cells were in agreement with the HER2 SISH results in 37.50% (12 out of 32 cases), 43.75% (14/32), and 43.75% (14/32) of the specimens, respectively (Table 1). Therefore, we decided to use the approach of analyzing a certain area ($40457.64 \mu\text{m}^2$) until FOVs with at least 1000 tumor cells were assessed. We did not exclude intervening normal structures while analyzing the FOVs. Subsequently, 555 breast cancer specimen slides were scanned and analyzed using the optimized image-based assessment and interpretation method, and the results were compared with those from the manual method.

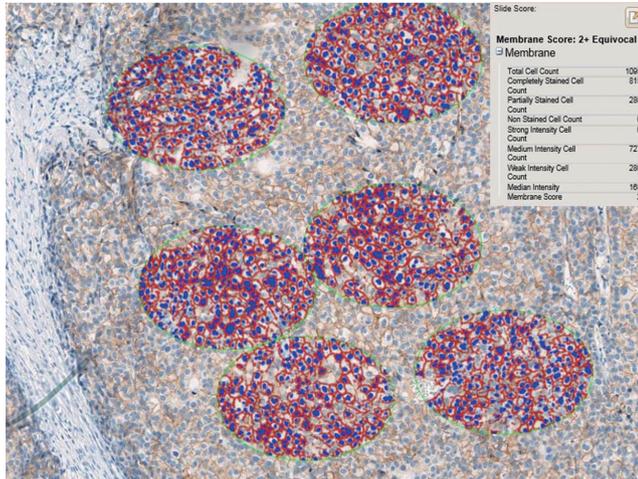


Fig. 1. Drawing of several ellipses, each with a constant area of 40457.64 μm², until 500, 1000, or 2000 cells were included in the focal view, was used to determine the optimal interpretation numbers (×200).

Table 1. Image Analysis Results for HER2 Equivocal (2+) Cases as Determined by a Manual Scoring Method in 2013

TC*	HER2 SISH	HER2 IHC			Total
		1+	2+	3+	
500	(-)	2	10	0	12
	(+)	0	10	10	20
1000	(-)	4	8	0	12
	(+)	0	10	10	20
2000	(-)	4	8	0	12
	(+)	0	10	10	20

HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; SISH, silver *in situ* hybridization.

*Total number of tumor cells evaluated.

Statistical analysis

For the comparison of the manual-scoring-based and image analysis-based IHC stain scoring results, Cohen’s kappa statistics were derived. Furthermore, sensitivity and specificity values were determined for both IHC methods by using the HER2 SISH results as a reference standard. Finally, McNemar’s test was performed on the SISH-positive and SISH-negative cases to determine whether there was a statistically significant difference in sensitivity and specificity values between the two methods.

RESULTS

Based on the manual scoring assessments of breast cancer specimens from 555 patients, 373 were categorized as score 0 (67.2%), 61 as score 1+ (11.0%), 46 as score 2+ (equivocal; 8.3%), and 75 as score 3+ (13.5%). During application of the image analysis method, the mean number of cells analyzed per case was 1107 (range, 1000 to 1328). The image analysis results showed that 282 specimens were categorized as score 0 (50.8%), 145 as score 1+ (26.1%), 51 as score 2+ (9.2%), and 77 as score 3+ (13.9%) (Fig. 2). The results obtained by manual scoring and image analysis had a concordance of 93.7% (Cohen’s kappa=0.830) (Table 2).

The HER2 SISH results revealed that 451 (81.2%) of the 555 cases were negative and 104 (18.7%) were positive. Among the HER2 IHC manual scoring results, all cases that were IHC positive (score 3+) were also SISH positive, and all cases that were IHC negative (scores 0 or 1+) were also SISH negative. Among the 46 manual-scored equivocal (score 2+) cases, 29

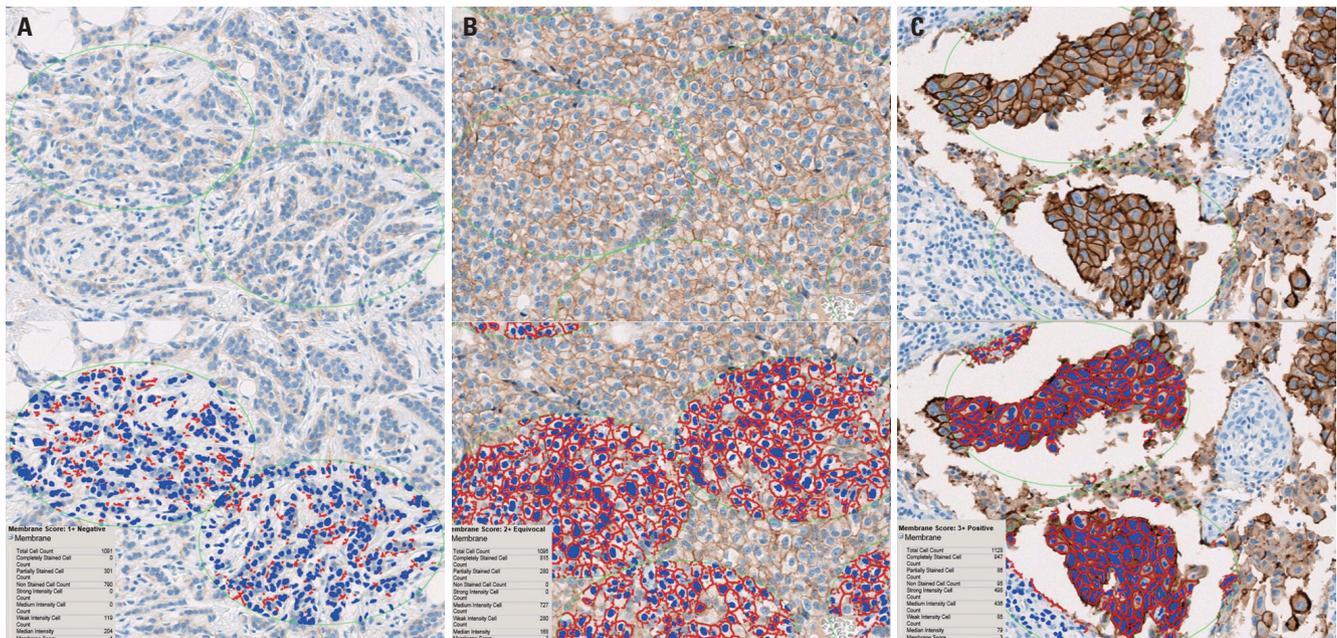


Fig. 2. Examples of image analysis scoring results. (A) HER2 IHC score 1, (B) HER2 IHC score 2, and (C) HER2 IHC score 3; ×200. HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry.

cases (63.0%) were SISH positive. Similarly, the image analysis-obtained HER2 IHC results showed that all cases that were IHC positive (score 3+) were also SISH positive, and all cases that were IHC negative (scores 0 or 1+) were SISH negative. Among the 51 equivocal (2+) cases identified by the image analysis method, 27 cases (52.9%) were SISH positive (Table 3).

The sensitivity, specificity, and accuracy of the manual scoring results were 72.1, 96.2, and 91.7%, respectively, while those derived by image analysis were 74.0, 94.7, and 90.8%, respectively. McNemar's test was applied to compare the HER2 IHC-staining results obtained by the manual scoring method to those obtained via the image analysis method. The test results showed non-significant differences between methods among both the SISH-positive cases ($p=0.688$) and the SISH-negative cases ($p=0.118$) (Table 4).

DISCUSSION

Typically, during HER2 image analysis in routine clinical labo-

Table 2. Comparison of HER2 IHC Scores from Manual and Image Analysis Methods for 555 Breast Cancer Specimens

Manual	Image analysis			Total
	0/1+	2+	3+	
0/1+	421	13	0	434
2+	6	31	9	46
3+	0	7	68	75
Total	427	51	77	555

HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry.

ratory practice, the pathologist selects and marks the FOVs, and the laboratory technician performs the HER2 image analysis in the marked FOVs. Subsequently, the pathologist reviews the results before their final approval. In this study, we sought to determine the minimal number of tumor cells to be evaluated during image analysis by analyzing 32 HER2 IHC manual-scored equivocal (2+) cases. We used the stamp function in the image analysis system to select same-sized areas repeatedly. A trained technician used the stamp function repeatedly to select areas with the most strongly stained tumor cells until 500, 1000, and 2000 tumor cells were included, regardless of the amount of tumor stroma components that was included. For technician training, a pathologist (K. Yim) taught the technician our newly developed method for HER2 image analysis, as well as basic histologic features of breast cancer and the HER2 IHC analysis guidelines for half a day. If it was difficult to see where the breast cancer was, the technician consulted and discussed with the pathologist and analyzed that case together. Finally, 50 cases were randomly selected to confirm that the analysis was correct. The results of HER2 image analysis with 500, 1000, and 2000 tumor cells were then compared with results of HER2 SISH. The results of HER2 image analysis agreed with those of HER2 SISH more closely when 1000 tumor cells were analyzed than when 500 tumor cells were analyzed, and there was no difference between the results when 1000 and 2000 tumor cells were analyzed (Table 1). On that basis, we applied the stamp function HER2 image analysis process to acquire at least 1000 tumor cells in areas with the most strongly stained tumor cells in 555 invasive breast cancer case specimens and compared the results with those of HER2 IHC man-

Table 3. Summary of HER2 IHC Scores from Manual and Image Analysis and HER2 SISH Results for 555 Breast Cancer Specimens

SISH	IHC result by manual method				IHC result by image analysis				Total (%)
	0	1+	2+	3+	0	1+	2+	3+	
(-)	373	61	17	0	282	175	24	0	451 (81.2)
(+)	0	0	29	75	0	0	27	77	104 (18.7)
Total (%)	373 (67.2)	61 (11.0)	46 (8.3)	75 (13.5)	282 (50.8)	175 (26.1)	51 (9.2)	77 (13.9)	555 (100)

HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; SISH, silver *in situ* hybridization.

Table 4. Comparison of HER2 Results Obtained by Manual and Image Analysis Methods for Predicting Positive (3+) or Negative (0, 1+) Results of HER2 IHC among HER2 SISH-Positive or SISH-Negative Patients

	Image analysis		Total	McNemar's test
	2+	3+		
SISH-positive cases				
Manual method	2+	25	4	0.688
	3+	2	73	
Total		27	77	104
0 or 1+				
SISH-negative cases				
Manual method	0 or 1+	423	11	0.118
	2+	4	13	
Total		427	24	451

HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; SISH, silver *in situ* hybridization.

ual scoring and HER2 SISH assays. The HER2 image analysis and HER2 manual scoring results showed a high level of agreement (Cohen's kappa=0.830) (Table 2). This supports the suggestion that, if a laboratory has a relatively high frequency of equivocal HER2 IHC results, integrating HER2 digital image analysis into the HER2 assessment routine of the laboratory can reduce the frequency of equivocal cases, which can lead to fewer HER2 ISH confirmatory tests (e.g., 34% by manual assessment vs. 10.1% by image analysis).⁷ Another HER2 digital image analysis study reported that not only would there be a reduction in the frequency of equivocal (2+) cases (16.2% by manual assessment vs. 10.4% by image analysis), but also that image analysis can improve intra-observer reproducibility ($\kappa=0.766$ by manual assessment vs. $\kappa=0.882$ by image analysis).⁸ However, in this study, the rate of equivocal cases was similar (8.3% by manual scoring vs. 9.2% by image analysis). Moreover, we did not detect any statistical difference in sensitivity or specificity between the results of HER2 manual scoring and those from image analysis (based on the McNemar's test result) (Table 4). Our laboratory performed HER2 IHC and SISH in all invasive breast cancers at the time of this study and performed quality control checks regularly, comparing both results. In addition, HER2 IHC manual scoring was prescreened by a resident and signed by a breast pathologist (A. Lee) or by fellows undergoing breast pathology training. Considering that the percentage of equivocal cases measured by image analysis was approximately 10%, this study would not have shown any remarkable benefit because the equivocal percentage in routine manual assessment results performed by our laboratory was already at 8%. In previous image analysis studies for HER2 IHC, all tumor cells included were analyzed, requiring time and effort in selecting all tumor areas.^{7,8} We searched for a convenient and relatively accurate method of reading a single whole slide. We expect that if we apply the image analysis method described in this study (i.e., 1) using a consistent-area stamp function and disregarding the amount of included stroma components, 2) evaluating at least 1000 tumor cells, and 3) analyzing the most strongly stained area) in a laboratory's diagnostic routine, the time and effort spent assessing HER2 status would be greatly reduced without lowering diagnostic accuracy. In addition, there were no significant method-based discrepancies between the results for HER2 negative (0/1+) and HER2 positive (3+) specimens. Since all the discrepancies observed in this study were in equivocal (2+) specimens, the use of the image analysis method is acceptable because all equivocal cases should be further examined by ISH, according to the HER2 interpretation guideline.¹

In conclusion, we have shown that HER2 results obtained via image analysis are similar to those obtained via the manual scoring method. Adoption of an image analysis approach would save time and effort in a clinical setting. Although further confirmation and external validation is required, we suggest examining the most strongly stained area, evaluating at least 1000

tumor cells, and disregarding the amount of stroma content.

ACKNOWLEDGEMENTS

This study was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (NRF-2017R1D-1A1B03034165).

This study was supported by Roche Diagnostics, US. The sponsor had no role in the study design, data collection and analysis, preparation of manuscript and decision to publish.

AUTHOR CONTRIBUTIONS

Conceptualization: Ahwon Lee. Funding acquisition: Ahwon Lee. Data curation: Kwangil Yim, Hong Sik Park, Dong Min Kim. Analysis and interpretation of data: Kwangil Yim, Ahwon Lee. Project administration and supervision: Youn Soo Lee, Ahwon Lee. Validation: Youn Soo Lee, Ahwon Lee. Writing—original draft: Kwangil Yim, Ahwon Lee. Writing—review & editing: Kwangil Yim, Ahwon Lee.

ORCID iDs

Kwangil Yim	https://orcid.org/0000-0001-8767-9033
Hong Sik Park	https://orcid.org/0000-0003-0888-0711
Dong Min Kim	https://orcid.org/0000-0001-7133-4647
Youn Soo Lee	https://orcid.org/0000-0002-1653-6315
Ahwon Lee	https://orcid.org/0000-0002-2523-9531

REFERENCES

1. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* 2013;31:3997-4013.
2. Kim S, Park HS, Kim JY, Ryu J, Park S, Kim SI. Comparisons of oncologic outcomes between triple-negative breast cancer (TNBC) and non-TNBC among patients treated with breast-conserving therapy. *Yonsei Med J* 2016;57:1192-8.
3. Goldenberg MM. Trastuzumab, a recombinant DNA-derived humanized monoclonal antibody, a novel agent for the treatment of metastatic breast cancer. *Clin Ther* 1999;21:309-18.
4. Hudis CA. Trastuzumab--mechanism of action and use in clinical practice. *N Engl J Med* 2007;357:39-51.
5. Ross JS, Fletcher JA, Bloom KJ, Linette GP, Stec J, Symmans WF, et al. Targeted therapy in breast cancer: the HER-2/neu gene and protein. *Mol Cell Proteomics* 2004;3:379-98.
6. Dobson L, Conway C, Hanley A, Johnson A, Costello S, O'Grady A, et al. Image analysis as an adjunct to manual HER-2 immunohistochemical review: a diagnostic tool to standardize interpretation. *Histopathology* 2010;57:27-38.
7. Helin HO, Tuominen VJ, Ylisen O, Helin HJ, Isola J. Free digital image analysis software helps to resolve equivocal scores in HER2 immunohistochemistry. *Virchows Arch* 2016;468:191-8.
8. Minot DM, Voss J, Rademacher S, Lwin T, Orsulak J, Caron B, et al. Image analysis of HER2 immunohistochemical staining. Reproducibility and concordance with fluorescence in situ hybridization of a laboratory-validated scoring technique. *Am J Clin Pathol* 2012;137:270-6.