

Tissue Thickness Interferes With the Estimation of the Immunohistochemical Intensity: Introduction of a Control System for Managing Tissue Thickness

Shinobu Masuda, MD, PhD,* Ryohei Suzuki, ME,† Yuriko Kitano, BAg,‡ Haruna Nishimaki, MD, PhD,* Hiroko Kobayashi, MD, PhD,* Yoko Nakanishi, PhD,* and Hideo Yokoi, BSc§

Context and Objective: The conversion of immunohistochemical (IHC) results from 3-dimensional tissue to a 2-dimensional visual image without considering tissue thickness poses a considerable risk of misleading IHC intensities. The present study aimed to clarify whether tissue thickness interferes with the estimation of IHC staining intensity and to introduce a control system to manage it.

Design: We prepared cell lines that are used as controls for human epidermal growth factor receptor 2 (HER2) IHC (MDA-MB-231, MDA-MB-175VII, MDA-MV-453, and SK-BR-3), a polyclonal antibody for HER2, an interferometry to measure the tissue thickness of formalin-fixed paraffin-embedded sections, a microscope with a Halogen or an LED light source, a complementary metal-oxide semiconductor camera in which the output signal can be corrected to $\gamma = 1$, and a program to estimate color elements (hue, saturation, and luminance). It was examined whether tissue thickness interferes with the experimental scoring systems and practical classification of the routine HER2 scoring system.

Results: A noncellular control was shown to be better than a cellular control for managing tissue thickness. The IHC intensity for HER2 was correlated with tissue thickness ($R^2 = 0.8094$), even under the less-standardized condition, but this correlation was better under the improved standardized condition using corrected $\gamma = 1$ ($R^2 = 0.9282$). Discrepancies in practical HER2 scores were increased in sections with thicknesses < 2 and $> 5 \mu\text{m}$. A control system to manage tissue thickness was introduced.

Conclusions: Tissue thickness interferes with the estimation of the IHC intensity of HER2 in both experimental and practical scoring systems. A control system for managing tissue thickness is essential to increase the benefits of IHC as a standardized assay for clinical applications.

Key Words: tissue thickness, immunohistochemistry, standardization, HER2

(*Appl Immunohistochem Mol Morphol* 2021;29:118–126)

Received for publication November 9, 2019; accepted March 23, 2020.

From the *Department of Pathology and Microbiology, Division of Oncologic Pathology, Nihon University School of Medicine; †Planning, Molecular Diagnostics; ‡Strategic Planning and Molecular Diagnostics, Nichirei Biosciences Inc., Tokyo; and §Quality Control, Production & Technology, Global Innovation Center, Nichirei Biosciences Inc., Saitama Prefecture, Japan.

The results of this study are produced by the project, “Research and Development of Navigation System for Cancer Diagnosis and Treatment,” commissioned from the Japan Agency for Medical Research and Development (AMED) in fiscal 2016, under the scheme of “Development of Medical Devices and Systems for Advanced Medical Services; Research and Development of the Advanced Medical System by use of Medical Information (16he0302728h0003).

The authors declare no conflict of interest.

Reprints: Shinobu Masuda, MD, PhD, Department of Pathology and Microbiology, Division of Oncologic Pathology, Nihon University School of Medicine, Ohmaguchikami-cho 30-1, Itabashi-ku, Tokyo 173-8610, Japan (e-mail: masuda.shinobu@nihon-u.ac.jp).

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal’s website, www.appliedimmunohist.com.

Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Immunohistochemistry (IHC) has been used for decades. After the introduction of the fundamental principles of immunofluorescence techniques and immunoperoxidase methods,^{1,2} the basic technical procedures of IHC have been further developed, that is, the peroxidase-antiperoxidase method,³ horseradish peroxidase-labeled antibodies, the avidin-biotin complex method,⁴ and the tyramine amplification system,⁵ all of which enhance the signal of the protein detected by the antigen-antibody reaction in the tissue section. Antigen retrieval techniques have contributed to unmasking specific proteins, even if only small amounts of protein are presented, especially in formalin-fixed paraffin-embedded (FFPE) tissues. Proteolytic digestions^{6,7} and the heat-induced antigen retrieval technique⁸ remarkably increase not only the sensitivity of the target antigen protein but also the specificity with lower background noise.

For breast and gastric cancers, the scoring scheme for the readout of the human epidermal growth factor receptor 2 (HER2) IHC results is assessed based on a 0 to 3+ scale, which includes the estimation of the intensity of the staining. *HER2* gene amplification increases protein production and recruitment at the cellular membrane of cancer cells. The membranous expression of HER2 is captured by IHC and threshold-based scoring has been developed, making it possible to predict the amplification of

HER2 gene and the overexpression of its protein in most breast cancer cases, albeit with variable diagnostic accuracy depending on the overall IHC protocol condition. The American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) guidelines⁹⁻¹¹ have improved the algorithm for the estimation of IHC and in situ hybridization to select patients who can benefit from anti-HER2 antibody therapy.

There are variable factors in the process of the IHC procedure that may affect the results, including the condition of fixation, type of primary antibody, condition of antigen retrieval, and kind of detection method.¹² Standardization of these variable factors has progressed through contributions of medical communities as follows: the provision of kits approved by agencies in each country, such as the US Food and Drug Administration (FDA) or Pharmaceuticals and Medical Devices Agency (PMDA) in Japan; the spread of automated machines for immunostaining¹³; the development of external quality assessment systems (EQA), such as the College of American Pathologists, UK National External Quality Assessment Service (UK-NEQAS), Nordic Immunohistochemical Quality Control (NordQC), The Royal College of Pathologists of Australasia (RCPAQAP), and Japan Pathology Quality Assurance System (JPQAS); and the recommendations of guidelines published by academic associations, nonprofit organizations, and commercial companies that provide reagents and kits.

Recommendations for thickness sections in the Practical Guide to Specimen Handling in Surgical Pathology¹⁴ are 4 to 5 μm , 1 to 3 μm , 2 to 3 μm , 6 to 15 μm , and 6 to 12 μm for routine paraffin, renal sections, bone marrow, nerve histochemical staining, and amyloid demonstrations, respectively. The type of paraffin wax varies by country; in Japan, hard paraffin wax is used, which may affect the recommendation for tissue thickness. In Japan, the recommended tissue thickness is 4 μm for HER IHC and 5 μm for HER2 fluorescence in situ hybridization. According to recommendations, tissues are sectioned at the preset micrometer of microtomes; however, it is known that tissue thickness is affected by air quality, temperature, and humidity, as well as the skill level of technicians. One of the most fundamental conditions of sections, tissue thickness, remains unmet needs to be addressed for the standardization of IHC.

This study was conducted to clarify how tissue thickness interferes with the estimation of the intensity of IHC and to discuss a control system to standardize tissue thickness in the routine practice of pathology.

MATERIALS AND METHODS

Preparation of a Cellular Control and IHC

Breast cancer cell lines, MDA-MB-231, MDA-MB-175VII, MDA-MB-453, and SK-BR-3 (American Type Culture Collection), were used to make blocks as controls for HER2 score 0, score 1+, score 2+, and score 3+, respectively. MDA-MB-231, MDA-MB-175VII, and MDA-MB-453 cells were cultured in Leibovitz's L-15 medium (Life Technologies, Carlsbad, CA), and SK-BR-3 cells were cultured in McCoy 5a medium (Life Technologies) containing 10% fetal bovine

serum (SAFC Biosciences, Lenexa, KS) and 1% streptomycin (Life Technologies).

IHC staining for HER2 was performed using the Histofine HER2 Kit (POLY) (Nichirei Biosciences Inc., Tokyo, Japan) as described briefly below. Deparaffinized FFPE tissue sections were incubated with primary antibody (anti-HER2 polyclonal antibody) for 30 minutes at room temperature and then incubated with secondary antibody (goat anti-rabbit immunoglobulin G polyclonal antibody labeled with a peroxidase polymer) (Nichirei Biosciences Inc.) for 30 minutes at room temperature. Cell lines with scores of 0 and 3+ were used as negative and positive controls, respectively.

Because commercially available antibodies for β -actin were only monoclonal, a polyclonal anti- β -actin antibody was produced (Immuno-Biological Laboratories Co., Gunma, Japan). A synthetic peptide consisting of the amino acids (2 to 11) of cytoplasmic β -actin (Accession AAA51567) was created. The selected site was specific for cytoplasmic β -actin, differentiated from γ -actin (Accession CAA27723) and α -actin [cardiac muscle (Accession AAH09978), skeletal muscle (Accession NP_001091), and smooth muscle (Accession AAH17554)]. Antigen peptide at a concentration of 100 $\mu\text{g}/0.5\text{ mL}$ with emulsion was used to immunize 3 rabbits every 2 weeks up to 8 times. After the fourth and sixth immunization, blood samples were examined to check the increase in antibody levels by enzyme-linked immunosorbent assay. After the eighth immunization, whole blood samples were collected. The specificity of the raised antibodies was confirmed by Western blotting as a single band at 43 kDa. Briefly, the IHC staining protocol for FFPE tissue sections was performed as follows: primary antibody was used at a concentration of 1 $\mu\text{g}/\text{mL}$. Polymer secondary antibody was used to be the identical condition as well as the detection of anti-HER2 polyclonal antibody.

Interferometry Measurement of Tissue Thickness

A 3-dimensional (D) nontouch surface shape measurement system (VertScan 2.0) (Mitsubishi Chemical Holdings Group, Ryoka System Inc., Tokyo, Japan) was used for the interferometry measurement of tissue thickness. Two methods were compared with measure tissue thickness, a scribed method, and a mean method. For the scribed method, FFPE tissue was scribed in a linear manner by a sharp needle at a noncentral site, and the difference between the base of the glass slide and the height of the FFPE tissue was measured as the thickness of the section. For the mean method, the difference between the base of the glass slide and the height of the FFPE tissue was measured at 4 sides, and the mean difference of the 4 sides was taken as the thickness of the section. FFPE tissue samples of preset thickness at 2, 3, 4, 5, and 6 μm (10 each) were compared between the 2 methods.

If the tissue is small and located at the center of an FFPE section, a scratch does not affect the tissue; however, if the tissue is large and occupies a major area of the FFPE section, the described method will damage the tissue. Therefore, we utilized the mean method to measure tissue thickness.

Ten each of FFPE tissue samples cut at 2, 3, 4, 5, and 6 μm were measured and used for IHC of HER2 or

β -actin using cellular and a noncellular control. To compare the 3 scoring systems, FFPE presets at 2, 3, 4, and 5 μ m (5 each) were measured and utilized.

A Microscope and a Camera to Capture Standardized Images

To capture standardized images of IHC, the conditions of the microscope (AxioImager Z2) (Carl Zeiss Microscopy GmbH, Königsallee, Germany) and camera (EOS-1DC modified to support linear imaging) (Canon Inc., Tokyo, Japan) should be maintained as constant as possible. One of the most influential conditions of a microscope is the kind of light source. LED light provides a more stable color temperature and distribution of the spectrum independent of output voltage compared with a Halogen lamp light source (Fig. 1).

For cameras, critical points must be considered, especially for color elements (hue, saturation, and luminance). The output signal of an image sensor operates in a linear fashion defined as $V_{Sout} = V_{Sin}^\gamma$ and set at $\gamma = 1.0$. Usually, the linear output is converted by an image processor to nonlinear

fashion defined as $V_{Pout} = V_{Pin}^\gamma$ and set at $\gamma = 0.45$ so as to compensate display nonlinearity of $\gamma = 2.2$ in advance. However, unless γ is corrected to be 1 from 0.45, the captured image data of the microscope becomes darker than that of the real observation. The camera (EOS-1DC modified to support linear imaging) has a function to capture images with $\gamma = 1$ and store them in JPEG format. This linear JPEG image data enables us to evaluate the captured image data quantitatively.

The hue and saturation were corrected by using a standardized color chart provided by Dai Nippon Printing Co. Ltd (Tokyo, Japan) and a multi-SDI/HDMI monitor (LV5382; Leader Electronics Corp., Yokohama, Japan).

A Program to Estimate Color Elements (Hue, Saturation, and Luminance)

To estimate color elements (hue, saturation, and luminance), a program was developed and supplied by Canon Inc., one of the coinvestigators of this project, which has functions as follows: a function of vector scope that shows the distribution of the color element (hue, saturation, and luminance) of an original still image (Supplementary

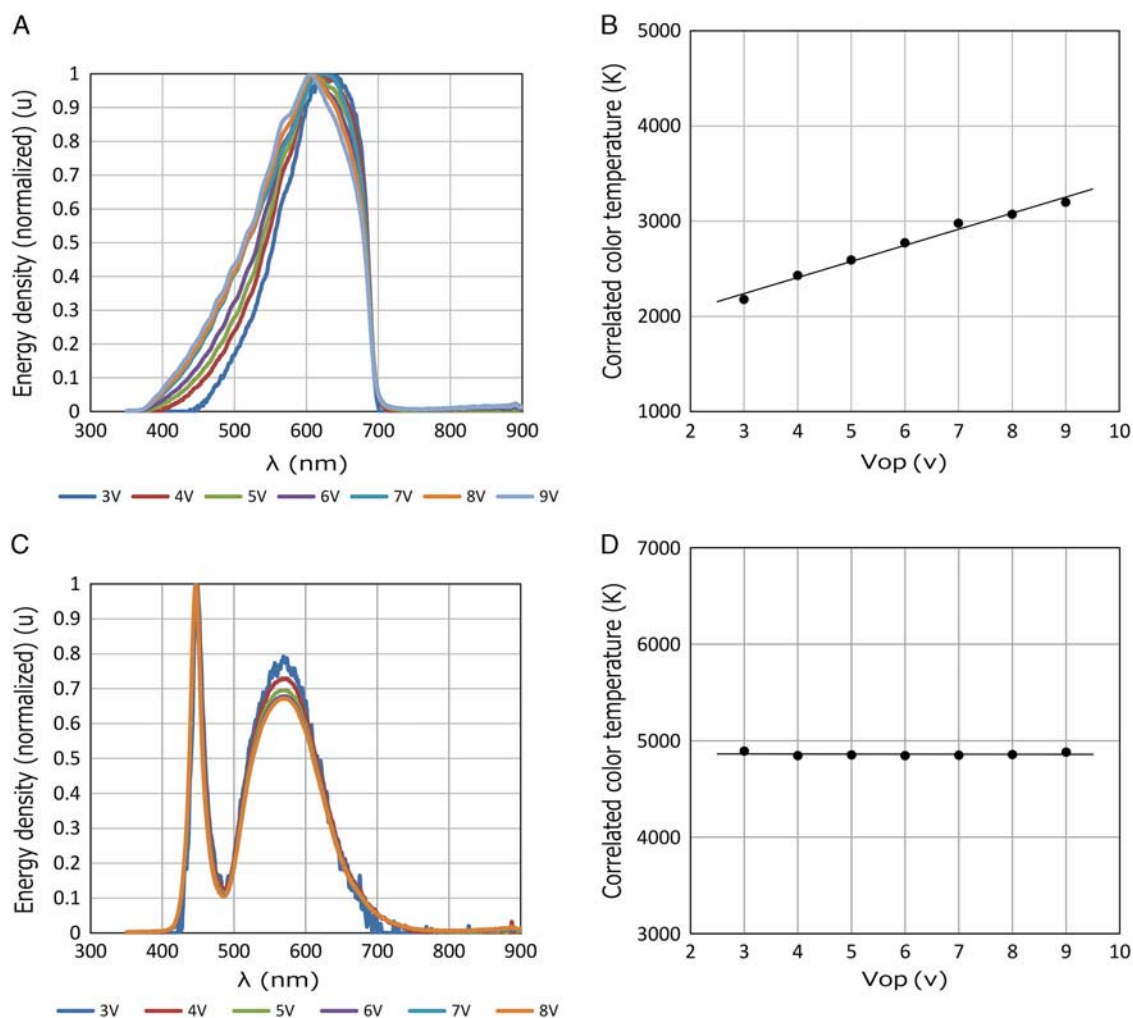


FIGURE 1. The influence of the light source of a microscope on color temperature. The spectra of energy density (A, C) and color temperature (B, D) at each output voltage (Vop) are shown using a Halogen lamp (A, B) and an LED light (C, D).

Figs. a, b, Supplemental Digital Content 1, <http://links.lww.com/AIMM/A259>); and a function to capture the color elements of a traced area and visualize the area with the same captured color elements in the whole original image (Supplementary Figs. c, d, Supplemental Digital Content 1, <http://links.lww.com/AIMM/A259>).

The manual estimation of IHC intensity to compare the preset thickness and luminance score (L score) was performed as follows: 10 each FFPE tissue sections of preset thickness at 2, 3, 4, 5, 6 μm were stained for HER2 or β -actin and estimated. Representative cells typically positive for HER2 score 3+ or β -actin were selected by a certified pathologist for each image and traced (Supplementary Fig. c, Supplemental Digital Content 1, <http://links.lww.com/AIMM/A259>). Then, we confirmed the positive area with the same color elements captured in the whole image area. The color elements and luminance of the captured area are shown in the vector scope with a histogram of luminance shown in each range of saturation. We can obtain data to calculate scores, that is, luminance, saturation, and pixel counts. This manual estimation of IHC intensity was used to analyze the correlation between tissue thickness and IHC intensity under less-standardized conditions.

For advanced analysis to compare 3 scoring systems, the estimated process was semiautomated based on the results obtained by the manual method. Using 10 FFPE tissue sections of preset thicknesses at 2, 3, 4, and 5 μm , stained for HER2 score 3+, 2+, and 1+, and the images were captured in a linear condition. The luminance and pixel counts were estimated within a hue of 70 to 170 degrees and saturation of 2 to 62, excluding the background $190 < R < 255$, $190 < G < 255$, and $190 < B < 255$. The definitions of the 3 scoring systems are described below.

Scoring Systems to Estimate of Intensity of IHC Staining

To estimate the intensity of IHC, 3 scoring systems were compared: an L score, a luminance with white balance score (L-WB score), and an intensity score (I score). The definition of each scoring method is as follows: L score = $\sum_{\text{saturation}} [(255 - \text{mean luminance}) / 255 \times (\text{brown pixels} / \text{total brown pixels}) \times 100]$, L-WB score = $[(\text{luminance of WB} - \text{mean of total luminance}) / (\text{luminance of WB}) \times 100]$, Mean of total luminance = $[\sum_{\text{saturation}} (\text{mean luminance} \times \text{brown pixels})] / \text{total brown pixels}$, I score \propto saturation/luminance \equiv mean of total saturation/mean of total luminance, Mean of total saturation = $[\sum_{\text{saturation}} (\text{mean saturation} \times \text{brown pixels})] / \text{total brown pixels}$ (Supplementary Table 1, Supplemental Digital Content 2, <http://links.lww.com/AIMM/A260>). To adjust the white balance, 2 methods were compared: an absolute white balance and a background white balance. For adjustment with the absolute white balance method, a glass slide covered by a glass cover without tissue sections was captured and estimated as white balance. The background white balance was captured in an area without a tissue in each slide.

Image Analysis to Determine Practical HER2 Score

To validate whether tissue thickness interferes with the practical classification of HER2 score, the same images used

for experimental estimation by the 3 scoring systems were analyzed. The images were submitted to a commercially available image analysis system (e-Memb 2; e-Path Co. Ltd, Kanagawa, Japan). A practical classification of HER2 score was determined for each image of FFPE (3 cell lines: score 1+, 2+, 3+; preset thickness: 2, 3, 4, 5 μm ; 5 sections for each condition) with measured tissue thickness.

Noncellular Control Materials to Manage Tissue Thickness in Routine Practice

In routine practice, it is difficult to measure the thickness of each FFPE section by interferometry, a time-consuming and effort-consuming method. To standardize tissue thickness for practical use, 2 control materials were compared: a cellular control and a noncellular control. As a cellular control, cell lines stained with β -actin, which was described in detail in the cell lines and IHC section, were used. As a noncellular control, a urethane form (Nissin Resin Co. Ltd, Kanagawa, Japan) was used, which was produced as follows. Briefly, the main agent was polyether polyol with a mixture of phthalocyanine blue (4.5 wt%), dimethylpolysiloxane as an exhibition powder (0.1 to 0.2 wt%), and H_2O as a foaming agent (0.1 to 2 wt%). The curing agent methylenebis (4, 1-phenylene) with non-halogen phosphoric acid ester was mixed with the main agent, stirred, and poured into the mold (100 \times 60 \times 30 mm). The urethane form was cut to an adequate size (20 \times 15 \times 5 mm), embedded in paraffin, and then hollowed out and transferred to a recipient block using a Tissue Microarrayer (KIN-1) (Azumaya Corporation, Tokyo, Japan).

Statistical Analysis

The correlation between tissue thickness and estimation of IHC intensity was shown in a scatterplot with a correlation line and coefficient of determination. Statistical analyses were performed using Microsoft Office Excel software (Microsoft Co., Redmond, WA).

RESULTS

Noncellular Control Was Better Than a Cellular Internal Control to Evaluate Tissue Thickness

As a control to manage tissue thickness, a cellular internal control stained by an internal protein (β -actin) and a noncellular control (a urethane form) were examined.

The cellular internal control showed a correlation between the L score and tissue thickness with a coefficient of determination value of 0.788 (Fig. 2A), even under the less-standardized condition using a captured image ($\gamma = 0.45$) with manual evaluation of IHC intensity. However, the correlation was considerably decreased, especially at higher thicknesses.

The noncellular control was examined under the improved standardized condition, that is, captured image corrected in a linear condition ($\gamma = 1$) with semiautomated estimation and absolute background. The correlation between the L score and tissue thickness was better, with a coefficient of determination of 0.8376 (Fig. 2B) than the cellular control, even when the thickness of sections was $> 5 \mu\text{m}$.

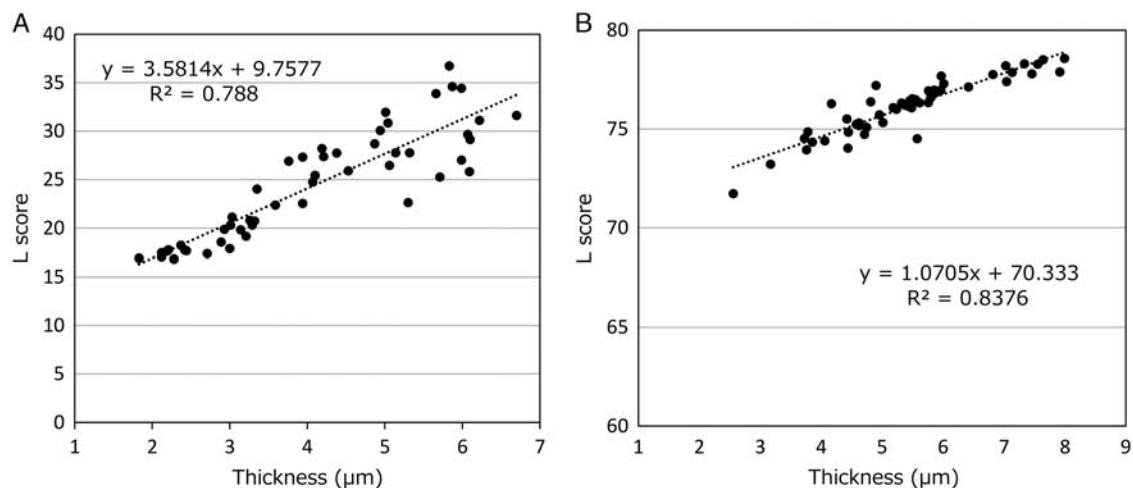


FIGURE 2. Correlation between tissue thickness and L score using cellular and noncellular control. The tissue thickness of a cellular internal control stained by an internal protein (β -actin) (A) and a noncellular control (an urethane form) (B) are compared with the immunohistochemical intensity score (L score). Tissue thickness was measured by interferometry for 10 of each preset at 2 to 6 μm . L score indicates luminance scoring method.

IHC Score and Tissue Thickness Increased in Parallel

Under the Less-standardized Condition

The visual images of samples of variable thickness provide different impacts on the intensity of staining, even if the sections were derived from the same sample (Fig. 3A). The tissue thickness was correlated with the L score with a coefficient of determination value of 0.8094 (Fig. 3B), even under the less-standardized condition using captured images ($\gamma=0.45$) with manual evaluation of IHC intensity and the L score.

Under the Improved Standardized Condition

The L score, L-WB score, and I score were compared (Fig. 4, Supplementary Table 2, Supplemental Digital Content 3, <http://links.lww.com/AIMM/A261>). The L score (Figs. 4A, D) and L-WB (Figs. 4B, E) score showed more uniformity with less standard variation than the I score (Figs. 4C, F). Regarding white balance, a background white balance (Figs. 4D–F) showed better results than an absolute white balance (Figs. 4A–C). For all of the scoring methods, HER2 score 2+ and score 3+ showed better results than score 1+ and score 0 (data not shown). There was a critical point to be noted for the L score. The L-WB score and I score have total luminance as a factor in the definition that reflects the luminance of the microscope when the image was captured, which enables us to correct the luminance and compare images. In contrast, the L score was defined by only the mean luminance of the captured image; therefore, a comparison between images with an approximate straight line was not available. Considering the results and correction possibility together, the L-WB score was most suitable for the estimation of IHC intensity.

The tissue thickness and IHC intensity were more highly correlated under the improved standardized condition using captured images corrected ($\gamma=1$) with semiautomated

image analysis and with the best scoring system (L-WB score with a background white balance), with a coefficient of determination value of 0.9282 (Fig. 4E).

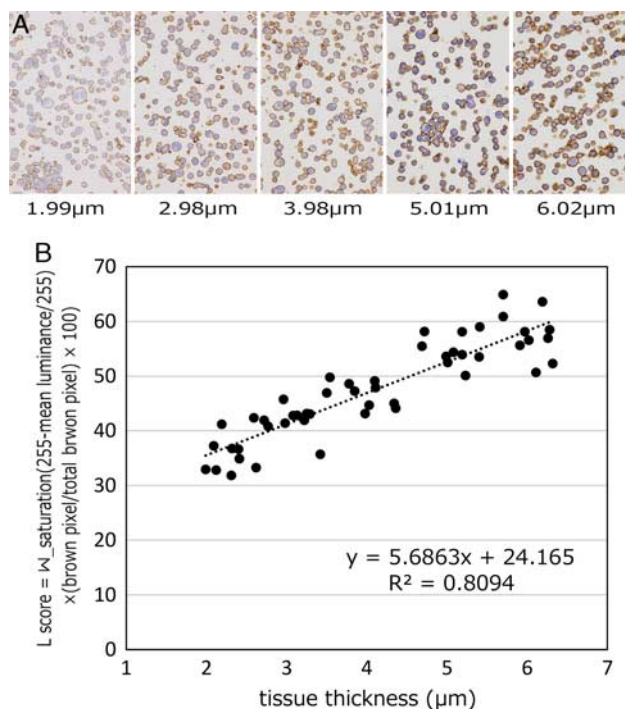



FIGURE 3. Effects of tissue thickness on the visual intensity of immunohistochemistry for HER2 under less-standardized conditions. The visual images of control (HER2 score 3+, SK-BR-3) with tissue thickness measured by the mean method using interferometry are shown (A). The correlation between tissue thickness measured by interferometry (10 of each preset at 2 to 6 μm) and immunohistochemical intensity score (L score) is shown (B). L score indicates luminance scoring method. 

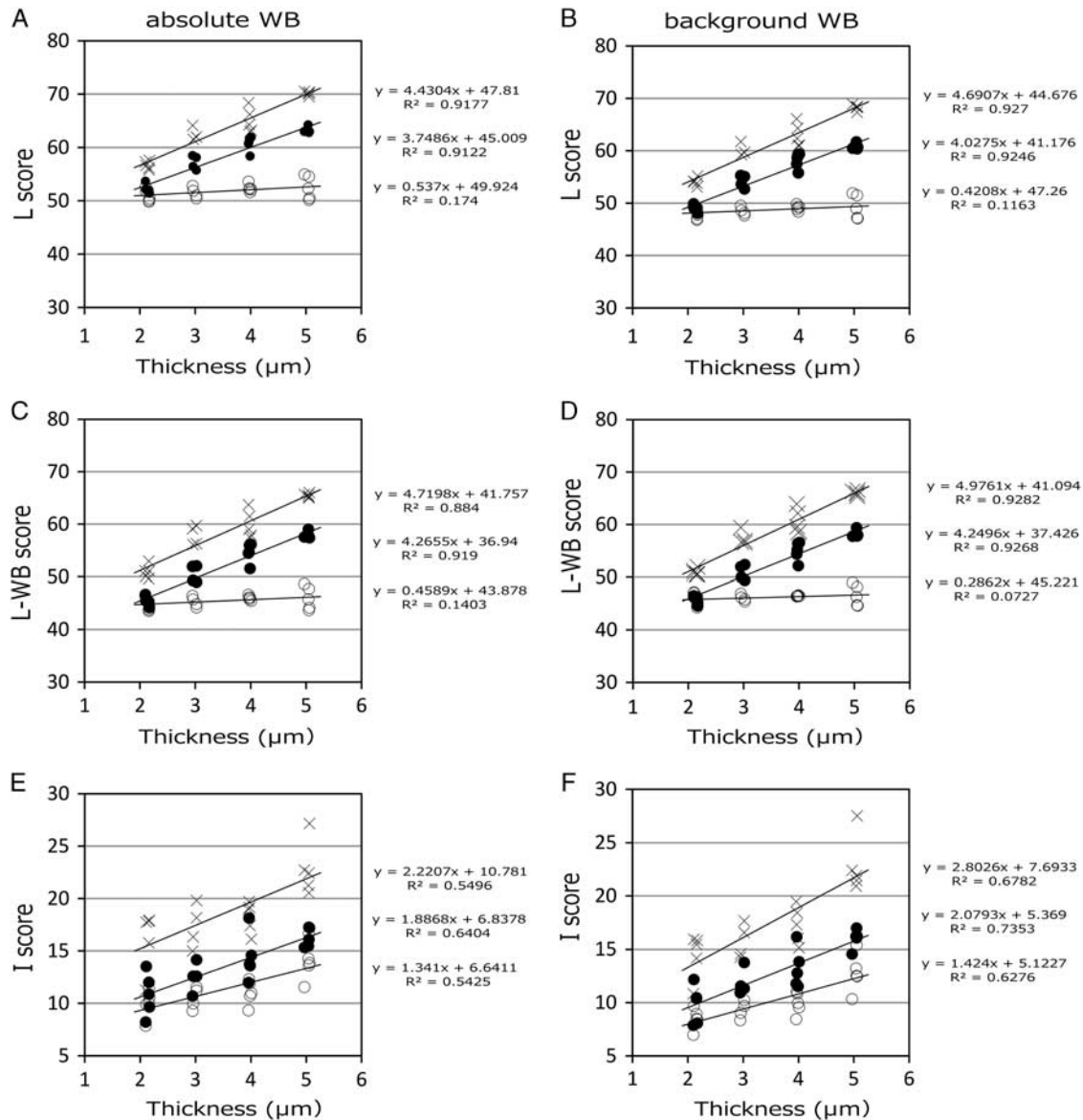


FIGURE 4. Tissue thickness and visual intensity of immunohistochemistry for HER2 under the improved standardized condition. Three scoring systems, L score (A, D), L-WB score (B, E), and I score (C, F), to estimate immunohistochemistry intensity and WB, absolute WB (A–C) and background WB (D–F) are compared with tissue thickness measured by interferometry (5 of each preset at 2 to 5 µm). I score indicates an intensity scoring method; L score, luminance scoring method; L-WB, luminance with white balance score; WB, white balance; ×, HER2 score 3+, SK-BR-3; ●, score 2+, MDA-MB-453; ○, score 1+, MDA-MB-175VII.

Tissue Thickness Interferes With the Practical Classification of HER2 Score

A practical classification of HER2 score, determined by an image analysis system according to ASCO/CAP guidelines (2018), is shown in Table 1. Control cell line sections for HER2 score 1+ (MDA-MB-175VII), score 2+ (MDA-MB-453), and score 3+ (SK-BR-3) were most adequately classified when tissue thickness was ~3 µm, given that scores 1+, 2+, and 3+ had discrepancies of 25%, 0%, and 0%, respectively. When tissue thickness was ~4 µm, the discrepancy was increased to score 2+ [100% according to ASCO/CAP guideline 2018; 60% according to the 30% rule (ASCO/CAP guideline 2003)],

although no discrepancy was detected for scores 1+ and 3+. The discrepancy of scoring classification was increased, especially when tissue thickness was ~2 µm for score 3+ (80%) and ~5 µm for score 2+ (100%).

Introduction of a Control System to Manage Tissue Thickness

A novel control system to manage tissue thickness is introduced in Figure 5. Controls in IHC are typically used to manage the staining procedures. However, as presented in this study and previous related studies, tissue thickness interferes with the estimation of IHC results through both

TABLE 1. The Tissue Thickness and Discrepancy of Practical HER2 Score Determined by Image Analyzing System

Measured Tissue Thickness (µm)		Practical HER2 Score Determined by Image Analyzing System																		
		2.1	2.11	2.16	2.16	2.17	2.95	2.96	3.02	3.02	3.96	3.97	3.98	3.99	4.01	4.97	5.04	5.05	5.05	5.06
cell line	MDA-MB-175VII (score 1)	0	0	1	1	ND	1	1	0	1	1	1	1	1	1	1	2	1	1	1
	MDA-MB-453 (score 2)	2	2	2	2	2	2	2	2	3 (2)*	3	3	3 (2)	3	3	3	3	3	3	3
	SK-BR3 (score 3)	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
		Percentage of Discrepancy (%)																		
Tissue Thickness (µm)		2					3					4					5			
cell line	MDA-MB-175VII (score 1)	50					25					0					20			
	MDA-MB-453 (score 2)	0					0					100 (65)					100			
	SK-BR3 (score 3)	80					0					0					0			

*30% rule by American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) guideline 2003.
 HER2 indicates human epidermal growth factor receptor 2.
 Shading cells show discrepant score between expected scores for cell lines and determined score by image analyzing system.

experimental scoring and practical scoring systems. A control system to manage tissue thickness is therefore required to accurately estimate IHC results.

An example of the proposed control system is shown in Figure 5A. A set of control material, including for both tissue thickness and staining procedures, is embedded in paraffin with sample tissue (Fig. 5A). In our study, FFPE sections were cut at the preset thickness, and the captured image of control material for tissue thickness was calibrated. If the thickness is adequate, the section is processed to IHC. Stained sections were further calibrated by controls for staining procedure. Only the sections with adequate thickness and staining are available for evaluation by pathologists (Fig. 5B).

DISCUSSION

The present study showed that tissue thickness interfered with the estimation of IHC intensity. The significance of tissue thickness for correct IHC evaluation was mentioned in an editorial article for the standardization of IHC.¹² The author noted that tissue thickness is one of the several overlooked variables that affect IHC results; tissue sections >5 µm can produce variation in stain intensity and the assessment of cytoplasmic and membrane staining is more complex than nuclear staining. An original article by Baker et al¹⁵ showed that the tissue thickness of the control slide is critical to manage the EQA system. Control slides should be more strictly standardized than sample slides since variation in tissue thickness of the control slides impedes constant evaluation based on the controls.^{15,16} A recently published report by McCampbell et al¹⁷ addresses the same issue of the effects of tissue thickness on IHC staining intensity. In their report, tumor markers, including Ki67, BCL6, CD7, and CK, were examined using whole-slide imaging. A study conducted by McCampbell and colleagues used interferometry to measure tissue thickness, calculated pixel percentage and was estimated by pathologists as gold standards similar to our study. The results presented in this study further support previous findings that tissue thickness interferes with IHC intensity estimation.

The impact of managing tissue thickness was validated for the practical classification of the HER2 score. In the present study system, tissue thickness of ~3 µm was deemed the best for accurate practical classification, and when tissue thickness was <3 or >4 µm, misclassification increased. The staining protocol utilized in this study was adopted for human breast cancer samples with approval

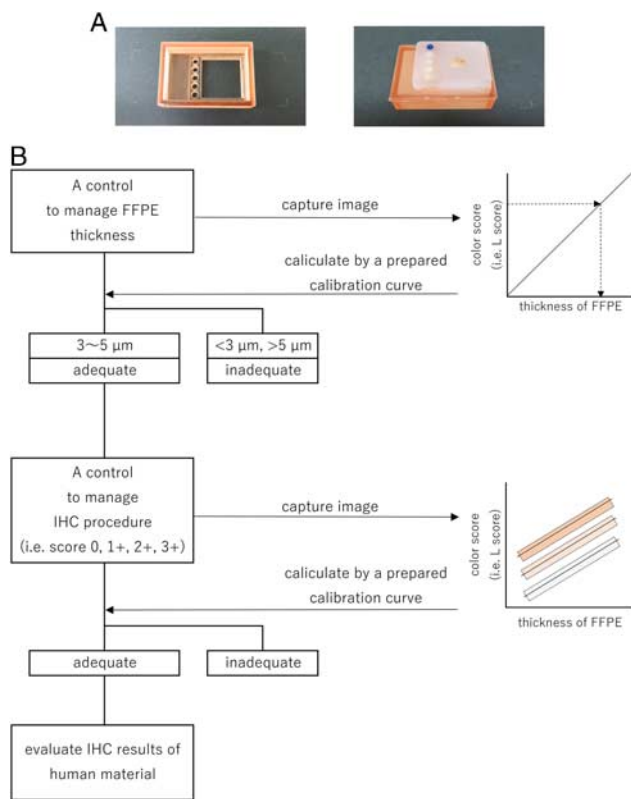


FIGURE 5. Introduction of a control system for managing tissue thickness in the routine practice of pathology. An example of a cassette for controls to manage tissue thickness and staining procedures is introduced (A). The thickness of the sections was calibrated using a standardized curve and control for tissue thickness. Only tissue sections with a range of thickness between 3 and 5 µm will be processed to IHC. IHC results are assessed by controls for staining procedures, and evaluated intensity will be further calibrated by the thickness (B). FFPE indicates formalin-fixed paraffin-embedded; IHC, immunohistochemistry.

from the PMDA in Japan. For control sections prepared using cultured cells lacking stroma, unlike human breast cancer, slight overstaining occurred. The kind of control that is best to manage staining procedures has yet to be addressed.

To manage tissue thickness, our first idea was to use cell lines stained for an internal protein, β -actin, as a control. Similar to other biological methodologies, such as enzyme-linked immunosorbent assays for protein or reverse transcriptional polymerase chain reactions for mRNA, we expected to be able to calibrate the basic contents of internal protein using cell lines stained for β -actin. However, this technique was found to be not suitable for the following reasons: first, as shown in the present results of the cellular control system (Fig. 2A), the correlation between tissue thickness and staining intensity was decreased, especially at the higher thickness. During the process to prepare a control section, cultured cells were exposed to extremely different conditions that changed their shape, such as being flattened on a cultured dish, placed in a medium, compressed by centrifugation, and embedded in the cell block. It is speculated that the thickness of compressed cells was not adequate to consistently reflect the actual tissue thickness. In contrast, the noncellular control was sufficiently homogenous and constant to reflect the actual tissue thickness, and it was also available for calibration. Second, it is theoretically difficult to obtain the ratio of aimed protein/internal protein considering the principles and protocols of IHC. During the process of IHC, antigen retrieval and polymer enhancement boost the signal visualization at various magnitudes; thus, the visualized intensity does not reflect the exact contents of the protein. Therefore, the present results showed that a noncellular control system was better than a cellular system for managing tissue thickness.

We introduced a control system for managing tissue thickness that is critically significant for medical care not only at present but also in the future when pathologic images will be more fundamentally utilized. Pathologic tissue sections are actually 3D but captured images from pathologic tissue are replaced with 2D images, and depth information is omitted during the conversion from 3D to 2D images. Tissue thickness enhances the intensity of staining as shown in this study and previous studies. Thus, IHC is incorrectly evaluated by computer-assisted image analysis systems. Furthermore, a vast amount of pathologic images are being accumulated as imaging data to construct artificial intelligence systems at present; however, these systems are learning from pathologic images without tissue thickness information, which is a fundamental factor. It is concerning that algorithms based on misclassified images will likely be led to the wrong evaluation systems. Thus, standardization of tissue thickness will be useful for evaluating IHC not only by the naked eye of pathologists but also by image analysis systems.

Our study had a few limitations. All experiments were based on cell blocks of cultured cell lines, and no human tissue samples were examined. Validation using different HER2 kits or different laboratory-developed tests should also be conducted. However, the present study aimed to clarify whether tissue thickness interferes the estimation of HER2 score under conditions where the

optimal system is controlled as strictly as possible (microscope, the captured image, and estimation system) because an optimal system has not yet been standardized for capturing images of IHC. Under the standardized optimal conditions, we confirmed the interference of tissue thickness as previously reported. On the basis of these results, we introduced a control system to manage tissue thickness. This concept should be applied to practical human samples using different approved kits and/or laboratory-developed testing kits in the future.

In conclusion, this study showed that tissue thickness interferes with the estimation of IHC intensity under the standardized optimal condition and proposed a control system for managing tissue thickness. Managing tissue thickness is of critical significance for medical care not only at present but also in the future when pathologic images will be more fundamentally utilized.

ACKNOWLEDGMENTS

The authors thank Yoshisuke Nakajima and Koji Meguro (NISSIN RESIN Co. Ltd) for their professional suggestions on urethane form; Ai Kakuhira and Yukari Hirotani (Division of Pathology, Nihon University School of Medicine) for their technical support; Dr Masato Takenaga (Nichirei Bioscience Inc.), Professor Michiie Sakamoto (Keio University School of Medicine) and Kumiko Nojima (Sakura Fintech Japan Co. Ltd) for their scientific discussions.

REFERENCES

1. Avrameas S. Coupling of enzymes to proteins with glutaraldehyde. Use of the conjugates for the detection of antigens and antibodies. *Immunochemistry*. 1969;6:43–52.
2. Nakane PK. Simultaneous localization of multiple tissue antigens using the peroxidase-labeled antibody method: a study on pituitary glands of the rat. *J Histochem Cytochem*. 1968;16:557–560.
3. Sternberger LA, Hardy PH Jr, Cuculis JJ, et al. The unlabeled antibody enzyme method of immunohistochemistry: preparation and properties of soluble antigen-antibody complex (horseradish peroxidase-antihorseradish peroxidase) and its use in identification of spirochetes. *J Histochem Cytochem*. 1970;18:315–333.
4. Hsu SM, Raine L, Fanger H. Use of avidin-biotin-peroxidase complex (ABC) in immunoperoxidase techniques: a comparison between ABC and unlabeled antibody (PAP) procedures. *J Histochem Cytochem*. 1981;29:577–580.
5. Adams JC. Biotin amplification of biotin and horseradish peroxidase signals in histochemical stains. *J Histochem Cytochem*. 1992;40:1457–1463.
6. Brozman M. Immunohistochemical analysis of formaldehyde- and trypsin- or pepsin-treated material. *Acta Histochem*. 1978;63:251–260.
7. Denk H, Radaszkiewicz T, Weirich E. Pronase pretreatment of tissue sections enhances sensitivity of the unlabelled antibody-enzyme (PAP) technique. *J Immunol Methods*. 1977;15:163–167.
8. Shi SR, Key ME, Kalra KL. Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections. *J Histochem Cytochem*. 1991;39:741–748.
9. Wolff AC, Hammond ME, Hicks DG, et al. American Society of Clinical Oncology, College of American Pathologists. 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. *Arch Pathol Lab Med*. 2014;138:241–256.
10. Wolff AC, Hammond ME, Schwartz JN, et al. American Society of Clinical Oncology/College of American Pathologists. American Society of Clinical Oncology/College of American Pathologists

- guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Arch Pathol Lab Med.* 2007;131:18–43.
11. Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *Arch Pathol Lab Med.* 2018;142:1364–1382.
 12. Leong AS, Leong TY. Standardization in immunohistology. *Methods Mol Biol.* 2011;724:37–68.
 13. Umemura S, Kurosumi M, Moriya T, et al. Recommendations for 'adequate evaluation of hormone receptors' a report of the task force of the Japanese Breast Cancer Society. *Oncol Rep.* 2010;24:299–304.
 14. Lott RT, Sheppard J, Santiago E, et al. *Practical Guide to Specimen Handling in Surgical Pathology.* College of American Pathologists; 2018:40–41.
 15. Barker C, Ibrahim M, Miller K, et al. Nondestructive quality control of HER2 control cell line sections: the use of interferometry for measuring section thickness and implications for HER2 interpretation on breast tissue. *Appl Immunohistochem Mol Morphol.* 2009;17:536–542.
 16. Leong AS. Quantitative immunohistology: tissue section thickness, another glitch in the path to standardization. *Appl Immunohistochem Mol Morphol.* 2009;17:465–469.
 17. McCampbell AS, Raghunathan V, Tom-Moy M, et al. Tissue thickness effects on immunohistochemical staining intensity of markers of cancer. *Appl Immunohistochem Mol Morphol.* 2019;27:345–355.