

## Comparative Analysis of Immunohistochemical Staining Intensity Determined by Light Microscopy, ImageJ and QuPath in Placental Hofbauer Cells

Katerina Cizkova<sup>1</sup>, Tereza Foltynkova<sup>1</sup>, Mariam Gachechiladze<sup>2</sup> and Zdenek Tauber<sup>1</sup>

<sup>1</sup>Department of Histology and Embryology, Faculty of Medicine and Dentistry, Palacky University, 779 00, Olomouc, Czech Republic and <sup>2</sup>Department of Clinical and Molecular Pathology & Laboratory of Molecular Pathology, Faculty of Medicine and Dentistry, Palacky University, 779 00, Olomouc, Czech Republic

Received November 19, 2020; accepted January 24, 2021; published online February 20, 2021

Software based analyses of immunohistochemical staining are designed for obtaining quantitative, reproducible, and objective data. However, often times only a certain type of positive cells or structures need to be quantified thus whole image analysis cannot be performed. Such an example is Hofbauer placental cells, which show positivity of some antigens together with trophoblast, but only Hofbauer cells represent the regions of interest (ROIs). Two independent observers evaluated the immunohistochemical staining intensity of Hofbauer cells in placenta samples stained for cytoplasmic antigens by ImageJ, QuPath and light microscopy. Thus, the precise manual determination of ROIs, i.e. Hofbauer cells, was necessary. We detected low inter-observer variability in staining intensity. Almost perfect agreement between observers was reached for ImageJ and QuPath whilst substantial agreement was reached for light microscopy evaluation. As for the comparison of ImageJ, QuPath and light microscopy, the agreement of all three methods (identical immunohistochemical intensity) was achieved for 38.1% samples. The almost perfect agreement of staining intensities was reached between ImageJ and QuPath, and moderate agreement for comparison of the light microscopy to both software. Software analyses are much more time-consuming, thus their utilization is at least questionable to evaluate ROIs with selection.

**Key words:** immunohistochemical staining evaluation, image analysis, inter-observer variability

### I. Introduction

Immunohistochemistry (IHC) is an effective, well-established and widely accepted method for localizing the expression of a specific proteins in tissues and it is used in both clinical and research practice. Antigens of interest in formalin-fixed paraffin-embedded tissues are detected by specific antibodies. The visualization of the antibody-antigen reaction is accomplished by using the relevant chromogen [7, 13, 21].

The stained sample slides are generally evaluated

under the light microscope by trained pathologist or researcher. Semi-quantitative scoring systems are widely used to convert subject perception of IHC-marker expression into (semi)quantitative data, which is then used for statistical analyses and establishing of the conclusions. The existing clinical scoring process is based on two characteristics: overall staining intensity and the proportion of tissue or cells stained. The overall score of the staining intensity typically has four categories: negative (0), weak (1), moderate (2), and strong (3). H-score, Allred-score, and Immunoreactive score are considered as a “gold standard” of combined scoring system in IHC data evaluation and presentation. All these scoring systems use different categories for the proportion of tissues or cells stained [1, 7].

Manual examination of histological slides under

\* Correspondence to: Zdenek Tauber, MD, PhD, Department of Histology and Embryology, Faculty of Medicine and Dentistry, Palacky University Olomouc, Hnevotinska 3, 779 00 Olomouc, Czech Republic. E-mail: zdenek.tauber@upol.cz

a light microscope is considered as time-consuming. Manual scores are qualitative (negative/positive) or semi-quantitative (negative/weak/moderate/strong) and subjective. Even when assigned by trained observers, the inter-observer variability is one of the greatest issues associated with examination of histological slides under the light microscope [1, 8, 12, 15]. Based on this, it is not surprising that there is an effort to use computer-aided or automated evaluation of IHC staining to produce quantitative, reproducible and objective data [1]. These software solutions are both, commercial and open-sources ones. Open-source software provides a collaborative option for image analysis and offers a more cost-effective option for those who have infrequent or educational use of image analysis, whereas commercial software provides more personalized image analysis choices [2]. Open-source software, such as QuPath [3] and IHC profiler plugin for ImageJ [21] are suitable solutions for evaluation of IHC staining [4, 10, 11].

The software evaluation could represent time-saving approach for evaluation of the intensity of immunostaining in whole image [3, 21]. However, often only a certain type of positive cells or structures in the tissue are intended for analysis. This situation can occur in both, research and diagnostics. For example, in placenta tissue, where trophoblast and Hofbauer cells both are IHC positive but only Hofbauer cells represent the regions of interest (ROIs) (Fig. 1). A similar situation can occur with the detection of various proteins in the kidneys (e. g. matrix metalloproteinases, ion handling proteins, PAX8, UGT1). These proteins can be expressed simultaneously by multiple structures, such as proximal and distal tubules, but expression in each should be assessed separately [5, 14, 19, 20]. The same occurs in diagnostic process, for example for vimentin, which is used as a marker of mesenchymal origin of tumours but it also stains endothelium, vascular smooth muscles, macrophages, fibroblast etc.

In this study, we used placental tissue sample as a model. The same set of IHC stained samples was evaluated by two independent observers by the light microscope, and two open-source software, ImageJ and QuPath to aim a comparison of the inter-observer variability as well as the agreement of the used methods.

## II. Material and methods

### *Immunohistochemical staining*

In total, 42 samples of formalin or methacarn-fixed, paraffin-embedded human placenta samples obtained from the archive of the Department of Histology and Embryology, Faculty of Medicine and Dentistry, Palacky University in Olomouc were used for the analysis. Immunostaining for 5 different cytoplasmic antigens (CYP2C8, CYP2C9, CYP2J2, IL-1 $\beta$ , IL-10) was performed. The following primary antibodies were used: rabbit polyclonal CYP2C8 (Proteintech Group, 6546-1-AP), rabbit polyclonal CYP2C9 (Abgent, AP7881c), both at dilution

1:50, mouse monoclonal CYP2J2 (Novus Biologicals, NBP2-46419) at dilution 1:200, rabbit polyclonal antibodies against IL-10 (Abcam; ab34843) at dilution 1:400 and IL-1 $\beta$  (Novus Biologicals, NBP1-19775) at dilution 1:100. The antibodies were diluted in Dako REAL™ Antibody Diluent (Dako).

The samples were stained according to standard indirect two step immunohistochemistry in 4  $\mu$ m thick paraffin sections. After deparaffinization and rehydration, the samples were pre-treated by incubation in 5% H<sub>2</sub>O<sub>2</sub> (20 min, RT), after which heat induced antigen retrieval in citric buffer pH 6 (120°C, 15 min, Histos), and incubation with ProteinBlock (30 min, RT) was performed. Then, the samples were incubated with selected primary antibody for 1 hr at the RT. The detection of the proteins was performed by EnVision™ Detection System, Peroxidase/DAB, Rabbit/Mouse (Dako). Tris buffer (pH 7.6) was used for washing between the various steps. Nuclei of all samples were counterstained with haematoxylin. The samples were then dehydrated and cover-slipped.

For image analysis, the RGB images of different 5 fields of vision with resolution 2040  $\times$  1536 pixels saved as .jpeg were obtained by light microscope Olympus BX40 equipped with Olympus DP71 camera at magnification 400 $\times$  for each sample.

### *Evaluation of staining intensity*

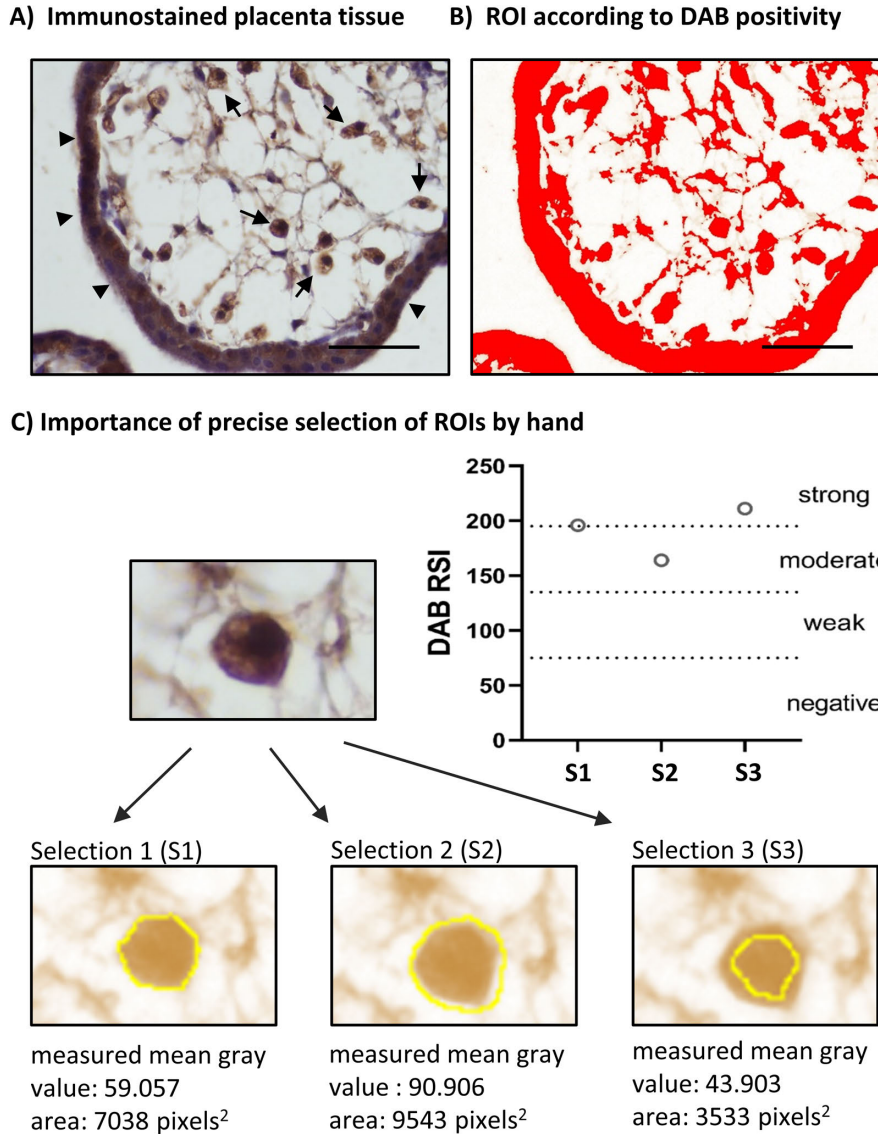
The evaluation of staining intensity of Hofbauer cells was performed by ImageJ and QuPath software and a light microscope by two experienced histologists. In addition, the time required to obtain the score for three samples was measured for both observers for all three methods.

### *Light microscopy*

The intensity of IHC staining was evaluated as: negative (0), weak (1), moderate (2), strong (3). The samples were evaluated twice in different times.

### *ImageJ*

The first step in analysis was colour deconvolution using IHC profiler plugin [21]. After that, the intensity of the immunostaining of Hofbauer cells was measured in deconvoluted DAB image. The Hofbauer cells (ROIs) were selected manually by observer 1 and 2 individually, and staining intensity was measured as “mean gray value” parameter. The average staining intensities for all measured cells from 5 fields of vision were counted for each sample. In ImageJ, the pixel intensity values for any colour range from 0 to 255, wherein 0 represents the darkest shade and 255 represents the lightest shade of the colour. Based on this, the staining intensities of the samples were divided into four groups (negative, weak, moderate, strong) according to thresholds established by creators of IHC profiler plugin [21] as follows: strong (3) for measured intensities ranking from 0 to 60, moderate (2) for intensities ranking from 61 to 120, weak (1) for intensities ranking from 121



**Fig. 1.** (A) Representative microphotograph of placenta tissue sample used in this study. The immunostained areas are represented by villous trophoblast (black arrows heads) and Hofbauer cells (black arrows), but the cells of the interest were only Hofbauer cells. Magnification 400×, Bar = 50 μm. (B) The selection of ROIs according to DAB positivity by threshold after colour deconvolution by IHC profiler plugin in ImageJ software. This approach is not possible, because all positive parts (trophoblast, Hofbauer cells) as well as background staining in stroma of the image are selected (red). Because only Hofbauer cells are cells of the interest for the analysis, ROIs must be determined manually. Bar = 50 μm. (C) Importance of precise selection of ROIs. The ROIs were selected manually after colour deconvolution in ImageJ software (yellow line). If ROI involved unstained surrounding of the cell, the measured intensity is lowered (Selection 2). The DAB staining intensities measured as “mean gray value” for Selections 1–3 are displayed in graph as reciprocal staining intensities (RSI;  $RSI = 255 - \text{mean gray value}$ ). Dotted lines represent thresholds for intensity categories.

to 180, and negative (0) for intensities higher than 181. In order to compare the distribution of measured staining intensities between ImageJ and QuPath, the values measured by ImageJ are displayed in graph as the reciprocal staining intensity (RSI), when  $RSI = 255 - \text{mean gray value}$  [13].

*QuPath*

The IHC staining intensity was measured as the optical density (OD). The staining vectors for all images were estimated automatically before OD measurement. The

Hofbauer cells (ROIs) were selected manually by observer 1 and 2 individually. To determine the intensity categories, the automatic thresholds were used. The samples with optical density (OD) lower than 0.2 were considered negative (0), for OD ranking from 0.2 to 0.4 weak (1), for OD ranking from 0.4 to 0.6 moderate, and for OD higher than 0.6 strong (3).

*Statistical analysis*

We evaluated the inter-observer variability as well as the agreement of the used methods. Statistical analysis was

performed using GraphPad Prism 8 software. All calculations were performed at the level of significance  $P > 0.05$ . The inter-observer variability for all three used methods was evaluated by weighted Kappa statistics. The overall score for each sample was determined as median of categories evaluated by observer 1 and observer 2, thus non-matching scores led to transient categories “negative/weak” (0.5), “weak/moderate” (1.5), and “moderate/strong” (2.5). The difference among ImageJ and QuPath software and light microscope was evaluated by Friedman test followed by Dunn’s multiple comparison. The strength of agreement between methods was evaluated by weighted Kappa statistics. For this calculation, the samples with inter-observer variability were excluded from the analysis. The measured time periods required for scoring were compared by one-way ANOVA followed by Dunn’s multiple comparison. The comparison of differences in time required for scoring between observers was performed by paired t-test.

### III. Results

#### *Inter-observer variability*

Scoring of immunohistochemical staining by human eye is known to be subjective. Moreover, the software evaluation of immunostaining intensities in our study involved “subjective” steps as well (i.e. precise selection of ROIs). Thus, we evaluated inter-observer variability for all used methods by weighted Kappa statistics. The strength of agreement between observer 1 and observer 2 was considered as “almost perfect” for ImageJ and QuPath software (weighted Kappa = 0.874 and 0.810 respectively) and “substantial” for light microscope evaluation (weighted Kappa = 0.741). The results are summarized in Figure 2. Moreover, there was no difference in time required to obtain scores for all three methods between observer 1 and observer 2 (paired t-test,  $P > 0.9999$ ).

#### *Comparison of ImageJ, QuPath and human eye evaluation*

The agreement of all three methods (i.e. identical IHC score) was achieved for 16/42 (38.1%). The representative microphotographs of samples with overall agreement and discrepancies in scoring among used methods are shown in Figure 3. The distribution of overall scores for all three used methods of evaluation is summarized in Figure 4. For non-matching scores obtained by observers, the transient categories “negative/weak” (0.5), “weak/moderate” (1.5), and “moderate/strong” (2.5) were used. The results showed that scores obtained by light microscope differed significantly from ImageJ and QuPath ( $P = 0.0010$  and  $P = 0.0056$  respectively; Friedman test followed by Dunn’s multiple comparison). The staining intensity was scored higher with light microscope. While both observers scored only one sample as strong positive using ImageJ and QuPath software, observer 1 scored 8/42 samples and observer 2 scored even 12/42 samples as strong positive.

The strength of agreement between methods was eval-

uated by Kappa statistics. The samples with inter-observer variability (falling into transient categories) were excluded from the analysis. The strength of agreement between ImageJ and QuPath software was considered as “almost perfect” (weighted Kappa = 0.945). The strength of agreement between light microscopy and software analysis showed “moderate agreement” with weighted Kappa = 0.446 for ImageJ and weighted Kappa = 0.527 for QuPath.

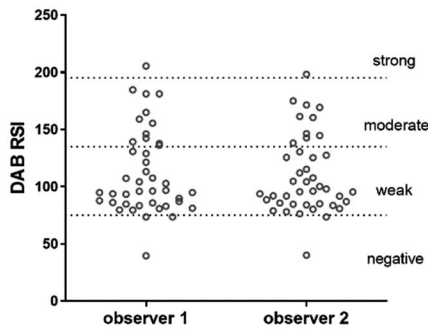
Because of the need to determine each Hofbauer cell as ROI, the software analyses were much more time-consuming in comparison to the evaluation by human eye. The time required for the scoring one sample significantly differed for the tested methods (one-way ANOVA,  $P > 0.0001$ ). The mean time for scoring of one sample was:  $10.8 \pm 1.9$  min for Image J,  $17.9 \pm 5.7$  min for QuPath, and  $2.1 \pm 0.4$  min for light microscopy ( $n = 6$  for each method).

### IV. Discussion

The aim of this study was to compare ImageJ, QuPath and human eye-based light microscopy evaluation of IHC staining of selected type of cells in tissue sections by two independent observers. In this study, the cells of interest were Hofbauer cells, placental macrophages localized in villous stroma, immunostained for different cytoplasmic antigens which also concurrently stain other parts of placenta tissue, such as trophoblast, the structure not intended for analysis. Thus, the obtained results presented here are independent on stained antigen, method of fixation and tissue processing, duration of storage and other known factors affecting IHC results.

Scoring of immunohistochemical staining under light microscope is known to be subjective and inter-observer variability represents the greatest issue in this approach. Inter-observer variability associated with human eye based assessment of IHC staining is influenced by the factors such as eye fatigue, complexity of data management following differential categorical scoring, quality of microscope, illumination of microscope and individual human vision limitations [6]. Many aspects of manual scoring can be impacted by visual traps (the phenomena in which the perceived image differs from objective reality) and cognitive traps (tendencies to think in a biased way that can lead to systematic errors or deviations from rational thinking) [1]. Contrary to this, digital image analysis via algorithms ensures that each section and each scoring event is viewed as an independent event, based on predefined metrics, and unaltered by the sections evaluated before or by adjacent cells [1]. The previous study performed by Jaraj *et al.* showed that the subjective assessment of intensity can be done with a high level of reproducibility while estimation of staining extent is less reliable [9]. Contrary to this, comparison of scores among 3 pathologist performed by Varghese *et al.* showed that scores may vary significantly [21]. In our study, both observers showed substantial agreement in scores. Moreover, the software evaluation of

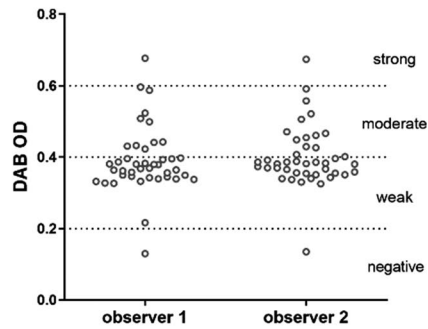
**A) ImageJ: Data distribution and inter-observer variability**



		observer 2				total
		negative	weak	moderate	strong	
observer 1	negative	2	1	0	0	3
	weak	0	26	0	0	26
	moderate	0	2	10	0	12
	strong	0	0	0	1	1
total		2	29	10	1	42

**weighted kappa = 0.874** (almost perfect agreement)

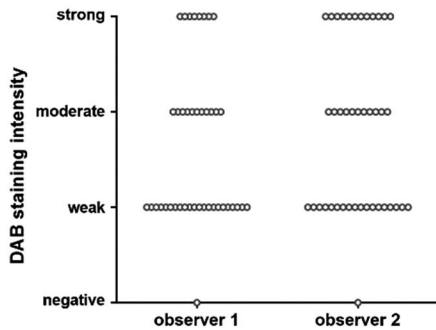
**B) QuPath: Data distribution and inter-observer variability**



		observer 2				total
		negative	weak	moderate	strong	
observer 1	negative	1	0	0	0	1
	weak	0	27	3	0	30
	moderate	0	1	9	0	10
	strong	0	0	0	1	1
total		1	28	12	1	42

**weighted kappa = 0.810** (almost perfect agreement)

**C) Light microscopy: Data distribution and inter-observer variability**



		observer 2				total
		negative	weak	moderate	strong	
observer 1	negative	1	0	0	0	1
	weak	0	17	5	0	22
	moderate	0	1	6	4	11
	strong	0	0	0	8	8
total		1	18	11	12	42

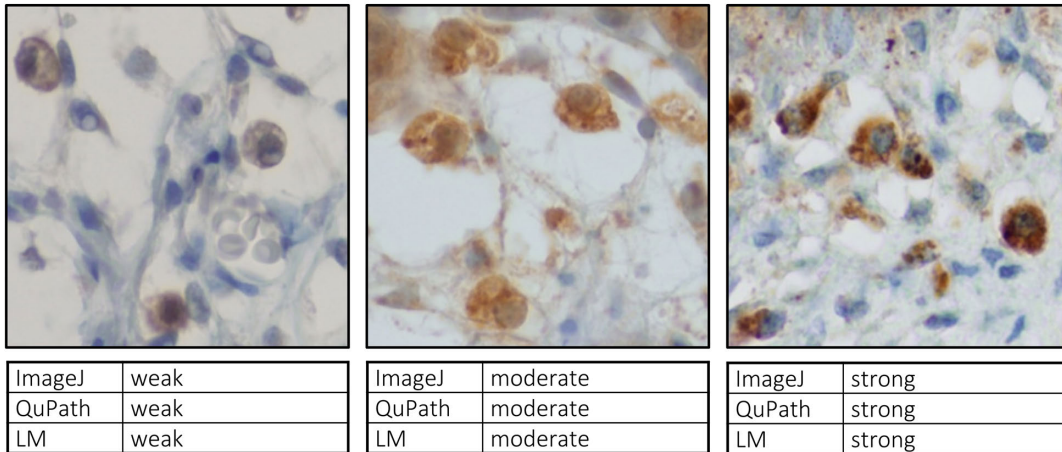
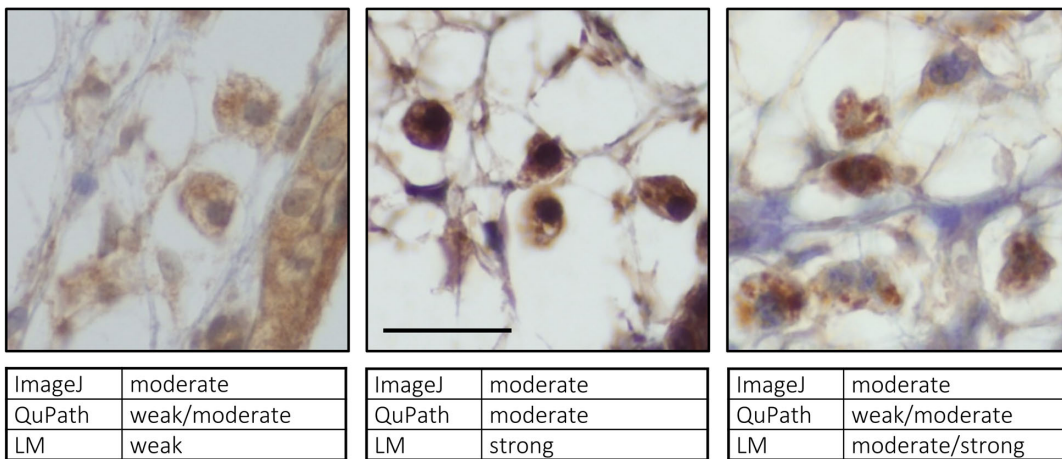
**weighted kappa = 0.741** (substantial agreement)

**Fig. 2.** Data distribution and inter-observer variability of scores obtained by (A) ImageJ software, (B) QuPath software, and (C) Light microscopy. The graphs show staining intensities measured by ImageJ and QuPath software and scores obtained by light microscope, each circle represents one sample (n = 42). In case of ImageJ measurements, darker areas have lower intensity values than lighter areas which is counterintuitive and the comparison of distribution of staining intensities between ImageJ and QuPath are cumbersome. To overcome this, the ImageJ intensities are displayed as RSI (reciprocal staining intensity). Dotted lines represent thresholds for intensity categories. The inter-observer variability was evaluated by Kappa statistics, indicating substantial or almost perfect agreement between observers for all three tested methods of immunostaining intensity evaluation.

immunostaining intensities in our study involved “subjective” steps as well. In this study, the precise manual determination of ROIs was necessary; not only determination of the cells of interest (Hofbauer cells) but also precise demarcation of these cells. Inaccurate demarcation involving negative cell surrounding distort the measured intensity value (Fig. 1). Our data showed almost perfect agreement for both, ImageJ and QuPath software.

Although the strength of agreement was almost perfect in scores assessed by ImageJ and QuPath software,

the comparison of these approaches with light microscopy showed only moderate agreement. The majority of non-matching cases were scored higher by light microscope assessment. Overall, our results are in agreement with previously published results by Ong *et al.* that majority of non-matching cases between light microscopy and computer evaluation were scored lower by computer. They also found the difference between cytoplasmic and nuclear antigens. Whereas light microscopy assessment scored higher than computer for nuclear antigens, they got opposite results

**A) agreement****B) discrepancy**

**Fig. 3.** Representative microphotographs showing (A) agreement, and (B) discrepancy and overall agreement (C) in immunostaining intensities obtained by ImageJ and QuPath software and light microscope. Agreement means that all three methods scored the sample in the same category. Discrepancy means that at least one of the methods scored the sample differently. All microphotographs have the same magnification (400 $\times$ ), Bar = 100  $\mu$ m. LM, light microscopy.

for cytoplasmic markers [16]. In this study, we examined only cytoplasmic markers thus our results differed from the study mentioned above at this point.

The human eye is more sensitive to higher intensity of IHC staining and it is least accurate at detecting differences under conditions of weak staining at which IHC is most linearly related to target antigen concentration. Machines are able to assess the colour intensity more accurately and provide the direct measurement of colour intensity from the hue, saturation and intensity index, whereas human eyes evaluate colour intensity in a more approximate manner [17, 18]. The immunostaining intensity of the cells is not uniform in whole sample. The lower scores obtained by software analysis in comparison to light microscopy may be caused by the fact, that results of software analysis are based on measurement of many cells in the sample and averaging of values. Whereas during assessment by

light microscopy, there should be the tendency to focus only on the cells with immunostaining of highest intensities which lead to higher scoring. Moreover, the localization of the studied antigen within cell cytoplasm should be taken into the consideration. If studied antigen is expressed strictly in certain region, for example around nucleus and not diffusely, and entire cytoplasm is evaluated uniformly, the machine may provide skewed (lower) results. In this case, the rest of the cytoplasm should be excluded from the evaluation. Although the software analysis approaches are designed for obtaining quantitative, reproducible and objective data [1], the quality of software image analysis approaches is highly influenced by a quality of tissue sections. This is due to the inability of most of the current automated image analysis systems to identify irregularities on a section that the human eye can ignore, such as artefacts, edge effect staining, folding of tissue and thickness of

**A) Agreement of scores between used methods**

		ImageJ				
		negative	weak	moderate	strong	total
QuPath	positivity					
	negative	1	0	0	0	1
	weak	1	23	0	0	24
	moderate	0	2	8	0	8
strong	0	0	0	1	1	
total		2	23	8	1	34

**weighted kappa = 0.945 (almost perfect agreement)**

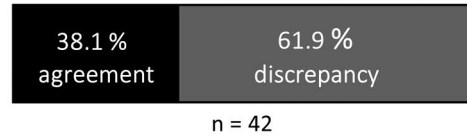
		ImageJ				
		negative	weak	moderate	strong	total
Light microscope	positivity					
	negative	1	0	0	0	1
	weak	0	15	1	0	16
	moderate	0	5	1	0	6
strong	0	1	6	1	8	
total		1	21	8	1	31

**weighted kappa = 0.446 (moderate agreement)**

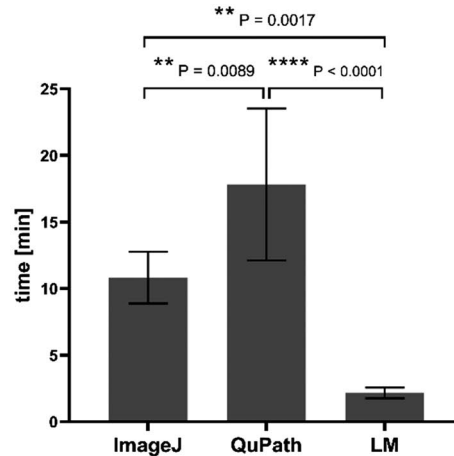
		QuPath				
		negative	weak	moderate	strong	total
Light microscope	positivity					
	negative	1	0	0	0	1
	weak	0	15	0	0	15
	moderate	0	5	1	0	6
strong	0	0	6	1	7	
total		1	20	7	1	29

**weighted kappa = 0.527 (moderate agreement)**

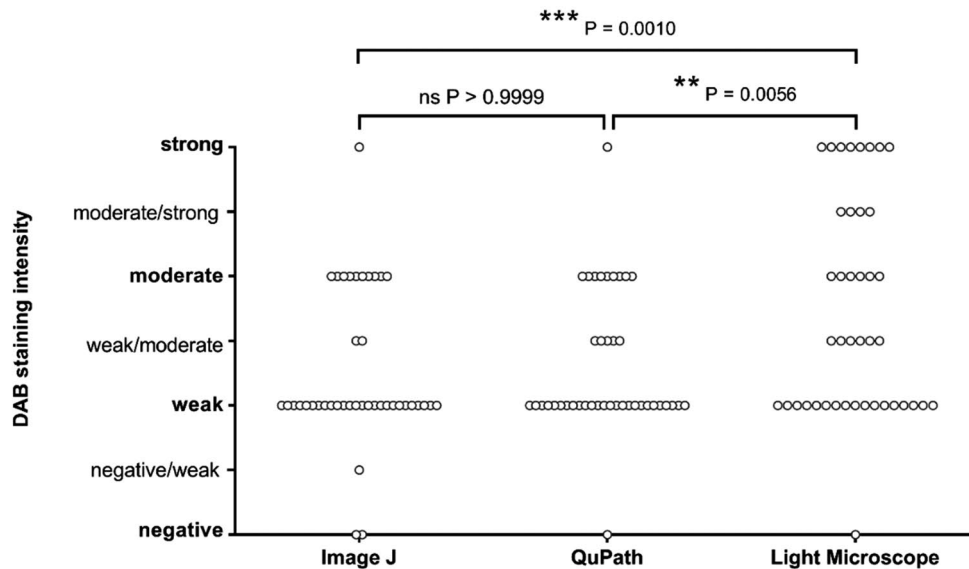
**B) Overall agreement of methods**



**C) Time required for scoring**



**D) Distribution of scores obtained by ImageJ, QuPath and Light microscope**



**Fig. 4.** Comparison of ImageJ and QuPath software and light microscope. **(A)** Agreement of scores between used methods. The table excludes samples wherein the inter-observer score did not match with each other. The results were evaluated by Kappa statistics. **(B)** Overall agreement. Agreement means that all three methods scored the sample in the same category. Discrepancy means that at least one of the methods scored the sample differently. **(C)** Time required for scoring. The time required for scoring was measured for three different samples for both observers (n = 6) and are displayed as mean ± SD. The results were evaluated by one-way ANOVA followed by Dunn’s multiple comparison. Statistically significant results are marked by asterisk (\*) directly in graph. \* P < 0.05, \*\* P < 0.01, \*\*\* P < 0.001. LM, light microscopy. **(D)** Distribution of scores obtained by ImageJ, QuPath and Light microscope. Each circle represents one sample (n = 42). Transient categories “negative/weak”, “weak/moderate”, and “moderate/strong” represents non-matching scores obtained by observers. The data were evaluated by Friedman test followed by Dunn’s multiple comparison. Statistically significant results are marked by asterisk (\*) directly in graph. \* P < 0.05, \*\* P < 0.01, \*\*\* P < 0.001.

the tissue section which may produce a false score [15].

Our data indicated that both tested software analyses required significantly longer time than light microscopy estimation. In case of whole image analysis, software analysis by ImageJ IHC profiler plugin could be a time-saving approach [21] but not in our case because of manual selection of ROIs (Hofbauer cells). Computer-assisted analysis takes less time to generate pathological scoring, it is objective and temporally linear regardless of the number of samples analysed [16]. To the contrary it is known, that time taken for visual scoring markedly increased with larger amounts of samples compared with computer scoring [16]. Fatigue is postulated as a potential source of error in visual interpretation of IHC stained tissue sections [22]. Because of software analysis used in our study required precise demarcation of ROIs, we suppose that fatigue could play role in all three methods.

Taken together, we compared ImageJ and QuPath software and human eye-based light microscopy evaluation of IHC staining of selected types of cells in tissue sections. The software analysis of IHC staining are designed for overcome inter-observer variability. We showed almost perfect agreement between two observers for ImageJ and QuPath IHC intensity scoring and substantial agreement for light microscopy estimation. The light microscopy scoring showed moderate agreement with ImageJ as well as QuPath method. In general, the majority of non-matching cases among tested methods were scored higher by light microscope assessment. The main disadvantage of software analyses in our study was a duration of time because precise selection of ROIs (in this study represented by individual Hofbauer cells) were needed. These approaches took significantly longer time than a light microscopy estimation evaluation. Thus in our opinion, the utilization of software analyses is at least questionable to evaluate ROIs with selection. In conclusion, it should be recommended to evaluate ROIs directly under light microscope by experienced histologists without questionable manual selection.

## V. Conflicts of Interest

Authors declare no conflict of interests.

## VI. Acknowledgments

This work was partly supported by IGA\_LF\_2020\_023.

## VII. Abbreviations

IHC, immunohistochemistry; ROI, region of interest; OD, optical density.

## VIII. References

1. Aeffner, F., Wilson, K., Martin, N. T., Black, J. C., Hendriks, C. L. L., Bolon, B., et al. (2017) The Gold Standard Paradox in Digital Image Analysis: Manual Versus Automated Scoring as Ground Truth. *Arch. Pathol. Lab. Med.* 141; 1267–1275.
2. Aeffner, F., Zarella, M. D., Buchbinder, N., Bui, M. M., Goodman, M. R., Hartman, D. J., et al. (2019) Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. *J. Pathol. Inform.* 10; 9–9.
3. Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., et al. (2017) QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* 7; 16878.
4. Bankhead, P., Fernández, J. A., McArt, D. G., Boyle, D. P., Li, G., Loughrey, M. B., et al. (2018) Integrated tumor identification and automated scoring minimizes pathologist involvement and provides new insights to key biomarkers in breast cancer. *Lab. Invest.* 98; 15–26.
5. Bellemare, J., Rouleau, M., Harvey, M., Popa, I., Pelletier, G., Têtu, B., et al. (2011) Immunohistochemical expression of conjugating UGT1A-derived isoforms in normal and tumoral drug-metabolizing tissues in humans. *J. Pathol.* 223; 425–435.
6. Conway, C., Dobson, L., O’Grady, A., Kay, E., Costello, S. and O’Shea, D. (2008) Virtual microscopy as an enabler of automated/quantitative assessment of protein expression in TMAs. *Histochem. Cell Biol.* 130; 447–463.
7. Fedchenko, N. and Reifenrath, J. (2014) Different approaches for interpretation and reporting of immunohistochemistry analysis results in the bone tissue—a review. *Diagn. Pathol.* 9; 221.
8. Fiore, C., Bailey, D., Conlon, N., Wu, X., Martin, N., Fiorentino, M., et al. (2012) Utility of multispectral imaging in automated quantitative scoring of immunohistochemistry. *J. Clin. Pathol.* 65; 496–502.
9. Jaraj, S. J., Camparo, P., Boyle, H., Germain, F., Nilsson, B., Petersson, F., et al. (2009) Intra- and interobserver reproducibility of interpretation of immunohistochemical stains of prostate cancer. *Virchows Arch.* 455; 375–381.
10. Loughrey, M. B., Bankhead, P., Coleman, H. G., Hagan, R. S., Craig, S., McCorry, A. M. B., et al. (2018) Validation of the systematic scoring of immunohistochemically stained tumour tissue microarrays using QuPath digital image analysis. *Histopathology* 73; 327–338.
11. Mane, D. R., Kale, A. D. and Belaldavar, C. (2017) Validation of immunoexpression of tenascin-C in oral precancerous and cancerous tissues using ImageJ analysis with novel immunohistochemistry profiler plugin: An immunohistochemical quantitative analysis. *J. Oral Maxillofac. Pathol.* 21; 211–217.
12. Melo, R. C. N., Raas, M. W. D., Palazzi, C., Neves, V. H., Malta, K. K. and Silva, T. P. (2020) Whole Slide Imaging and Its Applications to Histopathological Studies of Liver Disorders. *Front. Med. (Lausanne)* 6; 310–310.
13. Nguyen, D. (2013) Quantifying chromogen intensity in immunohistochemistry via reciprocal intensity. *Cancer InCytes* 2(1): e.
14. Ogbureke, K. U. E. and Fisher, L. W. (2005) Renal expression of SIBLING proteins and their partner matrix metalloproteinases (MMPs). *Kidney Int.* 68; 155–166.
15. O’Hurley, G., Sjöstedt, E., Rahman, A., Li, B., Kampf, C., Pontén, F., et al. (2014) Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers. *Mol. Oncol.* 8; 783–798.
16. Ong, C. W., Kim, L. G., Kong, H. H., Low, L. Y., Wang, T. T., Supriya, S., et al. (2010) Computer-assisted pathological



- immunohistochemistry scoring is more time-effective than conventional scoring, but provides no analytical advantage. *Histopathology* 56; 523–529.
17. Rimm, D. L. (2006) What brown cannot do for you. *Nat. Biotechnol.* 24; 914–916.
  18. Rizzardi, A. E., Johnson, A. T., Vogel, R. I., Pambuccian, S. E., Henriksen, J., Skubitz, A. P., *et al.* (2012) Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. *Diagn. Pathol.* 7; 42.
  19. Tong, G.-X., Yu, W. M., Beaubier, N. T., Weeden, E. M., Hamele-Bena, D., Mansukhani, M. M., *et al.* (2009) Expression of PAX8 in normal and neoplastic renal tissues: an immunohistochemical study. *Mod. Pathol.* 22; 1218–1227.
  20. van Raaij, S., van Swelm, R., Bouman, K., Cliteur, M., van den Heuvel, M. C., Pertijs, J., *et al.* (2018) Tubular iron deposition and iron handling proteins in human healthy kidney and chronic kidney disease. *Sci. Rep.* 8; 9353.
  21. Varghese, F., Bukhari, A. B., Malhotra, R. and De, A. (2014) IHC Profiler: an open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples. *PLoS One* 9; e96801.
  22. Weaver, D. L., Krag, D. N., Manna, E. A., Ashikaga, T., Harlow, S. P. and Bauer, K. D. (2003) Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer. *Mod. Pathol.* 16; 1159–1163.

---

This is an open access article distributed under the Creative Commons License (CC-BY-NC), which permits use, distribution and reproduction of the articles in any medium provided that the original work is properly cited and is not used for commercial purposes.

---