P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

Quantitative Histomorphometrical Analysis in Distinguishing Non-Hodgkin Malignant Lymphoma, Large Cell Type with Nasopharyngeal Carcinoma

Okky Husain[1], Hermin Aminah Usman[1], Anglita Yantisetiasti[1], Astri Handayani,[2] Bethy S. Hernowo[1]
[1]Department of Anatomic Pathology, Universitas Padjadjaran
[2]School of Electronical Engineering and Informatics, Institut Teknologi Bandung Bandung

Corresponding Author: Okky Husain, dr
Anatomical Pathology, Faculty of Medicine, Universitas Padjajaran
Email: okky.husain@gmail.com

Received : 08-11-2022
Accepted : 25-11-2022
Published: 31-01-2024

ABSTRACT

Background
Non-Hodgkin's Malignant Lymphomas (NHMLs) with large cell morphology have similar histopathological appearance with Nasopharyngeal Carcinoma, undifferentiated type therefore ancillary examination such as Immunohistochemistry (IHC) are required. Quantitative histomorphometrical analysis able to identify subtle detail. The aim of this study is to measure the accuracy of quantitative histomorphometrical analysis to distinguishing NHML, large cell type with Nasopharyngeal Carcinoma.

Methods
Up to 33 Hematoxylin and Eosin (HE) slides from NHML, large cell type, and 27 cases of Nasopharyngeal Carcinoma were photographed under 400 times magnification with a total of 7131 images. Background, nuclei, and extra-nuclear components between cells are segmented as mathematical objects. Objects' properties were measured and represented as distribution factors (minimum, quartile 1, median, quartile 3, maximum, range, and interquartile range). Significant factors that can differentiate the study groups are measured with a fisher-exact statistical test. Models to predict NHML, large cell type from nasopharyngeal carcinoma were formulated with different numbers of factors. The HE diagnosis accuracy compared to IHC was measured

Results
A sum of 7131 with 3893 from NHMLs, large cell type, and 3238 images of nasopharyngeal carcinoma were collected. After features selection, up to 76 properties factors are selected. Models' accuracy are ranging from 72.67% with a single factor and up to 89.9% with the whole selected factors.

Summary
Models' accuracy are higher (72.6-89.9%) compared to HE primary diagnosis (60%) in distinguishing NHML, large cell againts nasopharyngeal carcinoma. Quantitative histomorphometry can assist pathologist for early screening to distinguishing NHML, large cell type with Nasopharyngeal Carcinoma.

Keywords: Computational pathology, digital pathology, histomorphometry, lymphoma, malignant, nasopharyngeal carcinoma

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

INTRODUCTION

Malignant lymphoma is a primary malignancy of the lymph node system (KGB) and extranodal lymphoid tissue derived from lymphocyte cells. More than 95% of Malignant Lymphoma cases are Non-Hodgkin Malignant Lymphoma (NHML) with the remainder being Hodgkin's Lymphoma.[1] According to data from the Global Cancer Observatory (GLOBOCAN), NHML is the 7th most common and 9th most deadly cancer in Indonesia.[2] The most common cases of NHML in developing countries like Indonesia is Diffuse Large-B cell Lymphoma (DLBCL) which has a large neoplastic cell appearance.[3,4] The morphology of DLBCL can resemble other undifferentiated tumor cells.[5]

Nasopharyngeal carcinoma is the 5th most common and most lethal malignancy in Indonesia.[2] Nasopharyngeal carcinoma has several subtypes according to histopathological morphology. The morphological subtype of Non-Keratinizing Squamous Cell Carcinoma (NKSCC) is far more common than other subtypes of nasopharyngeal carcinoma.[6] Nasopharyngeal carcinoma and large cell NHML are often differentially diagnosed because of their similar microscopic appearance and both cancers are frequently encountered in Indonesia.

Immunohistochemical (IHC) examination functions to detect specific proteins in tissues. IHC examination can confirm a tumor as a carcinoma or as a malignant lymphoma.[5] Cytokeratin proteins are commonly used to confirm a tumor as a carcinoma, while malignant lymphoma is confirmed using a panel consisting of CD3, CD20, CD45, and Ki67.[7,8] the IHC examinations cannot yet be performed in all anatomical pathology laboratories. In addition, examination costs can be rise dramatically because the IHC examinations are often used as an examination panel with a combination of several proteins. It is necessary to optimize histopathological examination which is easily accessible and inexpensive.

Quantitative Histomorphometry is a computer-aided image analysis of digital histopathological microphotograph to identify specific pattern of diseases and predict their biological features. Submorphological features that too subtle for human eyes could be identified and analyzed by applying quantitative histomorphometry.[9,10] As access to information technology is getting cheaper and wider, the usage of it could be a solution to problem encountered above. In this study, authors are using quantitative histomorphometrical analysis of pathologists selected relevant region of interest to final diagnosis.

METHODS

Two different cases groups of NHML from head and neck region and nasopharingeal carcinoma were collected. Cases were from Hasan Sadikin General Hospital IHC laboratorium archives from year 2021. Cases were immunophenotypically confirmed. The clinical infomatrion were age, sex, the origin of specimens, and HE diagnosis prior to immunophenotyping.

The HE slides are collected for confirming the diffuse distribution of large neoplastic cell morphology. The carcinoma cases that included in the study has to shows no identifiable mature epithelial differentiation. Slides with poor quality will be optimized. If, after optimization, the quality of slides showed no improvement, the cases were excluded. Cases with identifiable differentiation therefore unneccesary to be differentially diagnosed were eliminated from the series.

Digital image acquization using Olympus CX3 trinocular microscope wth 0.5 magnification camera mount. Image acquizations are performed in selected region of interest that relevant to final diagnosis. Regions that were considered irrelevant are those dominated by necrosis, normal epithelial cells, normal connective tissue, and/or background and were not acquired. Whole slide images are not assembled in this study.

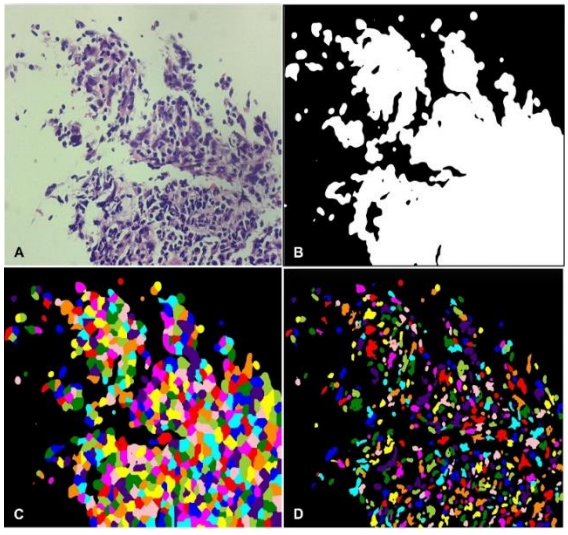

Figure 1. The image segmentation process. A. The micrograph from HE-stained B. White region represent tissue meanwhile black represent background C. Cellular segmentation. D. Labeled nuclear segmentation.

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

Global properties are measured as in table 1. The range of saturation and intensity are regularized with the maximum of value as 100%. Image segmentation are performed sequentially with the first step by segmenting brightfield background from tissue, next between nuclei and extra-nuclear component, and finally the extra-nuclear components between cells. The selected segmentation in our study is watershed segmentation algorithm. The segmentation 1. Histopathological properties for segmented regions are measured using *regioprop* function from *measure* module in Scikit-image.[11] The measured histopathology property are presented in Table 1.

Table 1. Measured histopathological properties.

Properties	Unit	Definition	Relevant Histopathological Features
Whole image properties			
Global intensity	%	The intensity values of the whole image	Overstaining/ Understaining
Global saturation	%	The saturation values of the whole image	
Cellularity	Cell/mm2	Amount of cells in a unit of area	Hypercellularity/ Hypocellularity
Cellular Properties			
Minimum cellular diameter	µm	The smallest diameter from the cellular margin	
Maximum cellular diameter	µm	The largest diameter from the cellular margin	Cellular Size
Cellular area	µm2	The area of the cell including cytoplasm and nucleus	
Cellular minimum to maximum diameter ratio	%	The smallest diameter are divided with the largest diameter of the cellular margin	
Mathematical eccentricity	%	The degree between ellips and perfect circle (0 is a perfect circle meanwhile the higher the value, the more ellips it is)	Round, oval, to spindle shape
Cellular extent	%	The ratio between celullar area and its bounding box area	
Nuclear–cytoplasmic (area) ratio	%	The area of nucleus divided by the cytoplasmic area of a cell	Nuclear–cytoplasmic ratio
Nuclear–cytoplasmic (diamater) ratio	%	The minimum diameter of nucleus divided by the minimum cytoplasmic diameter of a cell	
Cytoplasmic properties			
Cytoplasmic area	µm2	Cellular area subtracted by nuclear area	Cytoplasm amount
Cytoplasmic intensity	%	The value from cytoplasmic grayscale image	Clear cytoplasm
Nuclear properties			
Minimum nuclear diameter	µm	The smallest length between a point of nuclear margin with another margin	
Maximum nuclear diameter	µm	The largest length between a point of nuclear margin with another margin	Nuclear size
Nuclear area	µm2	Nuclear area	
Nuclear minimum to maximum diameter ratio	%	The minimum nuclear diameter divided by the maximum nuclear diameter	
Nuclear mathematical eccentricity	%	The degree between ellips and perfect circle (0 is a perfect circle meanwhile the higher the value, the more ellips it is)	Round, oval, to spindle shape
Nuclear Extent	%	The ratio between nuclear area and its bounding box area	
Nuclear Intensity	%	Mean intensitity value from nuclear grayscale image	Hyperchromaticity
Nuclear Membrane irregularity	%	4 pi times area divided by perimeter squared	Nuclear membrane irregularity
Nuclear Solidity	%	nuclear area divided by sum of nuclear area and f nuclear filled area	Vesicular chromatin

The distribution of each property is represented in a summary of 5 numbers representing the maximum value, 1st quartile, median, 3rd quartile and maximum value. Two other numbers are added, namely the range and interquartile range (IQR) of the property distribution. Outlier values will be stratified to be not lower than 1.5 IQR below 1st quartile and not higher than 1.5 IQR above 3rd quartile. In this study, each number is then called a factor.

Binary logistic regression measure-ments were carried out on each factor individually. Factors that did not show a significant difference between the two case groups were eliminated. Significance was assessed using the Fisher's exact test. Factors that pass the selection in the previous stage are then stratified based on the magnitude of their influence on the accuracy of each combination of factors. Stratification was performed using lasso regression.[12]

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

A binary logistic regression model was formulated from a combination of selected factors. The combination of factors used in this study is a combination of 5 factors, 10 factors, 20 factors, 50 factors, and all factors. The results of binary logistic regression predictions are displayed in the form of a 2x2 contingency table. Sensitivity was measured as the ratio of correctly predicted large cell NHML (true positive) compared to all cases confirmed as large cell NHML. Specificity was measured as the predicted ratio of non-NHML large cell and immunophenotypically confirmed carcinoma (true negative) to all carcinoma cases. Accuracy is the total of true positive and true negative cases compared to all cases. A fisher's exact statistical test was performed on the contingency table to measure the significance of the sensitivity, specificity and accuracy of the regression. Each factor coefficient of a combination of 5 factors, 10 factors, and 20 factors is analyzed and compared with the theoretical basis and other research.

RESULTS

Collected as many as 65 cases with 29 cases were nasopharyngeal carcinoma and 36 cases were large cell NHML. All cases of carcinoma showed cytokeratin immuno-expression. NHML cases showed immuno-expression of one of the lymphoma markers, namely CD20, CD45, or CD3. One case was classified as NHML after showing CD79a positivity. The other IHC panels are shown in Table 2. Paraffin blocks and preparations were found in all cases.

A total of 5 cases were eliminated, two cases were well differentiated carcinoma and metastatic lobular breast carcinoma which were examined for lymphoma and cytokeratin panel IHC. Three cases from the large cell NMHL group located outside the head and neck namely; breast folds, knees, and bone marrow. Other clinicopathological features are shown in Table 2.

Tabel 2. The clinicopathological features of included cases.

	NHML, Large Cell	Nasopharyngeal Carcinoma	P-value
Numbers (n)	33	27	
Sex (Male/Female)	17/16	21/6	0.04
Age: Mean (SD)	57.3 years (SD 9.5 years)	49.5 years (SD 7.9 years)	<0.01
HE primary diagnosis			
NKSCC, undifferentiated	9 (27%)	21 (78%)	
Unspecified Inflammation	4 (12%)	6 (22%)	
NHML, Large Cell	15 (46%)	-	
Hodgkin Lymphoma	5 (15%)	-	
Differential (secondary) diagnosis			
No differential diagnosis	3 (9%)	8 (30%)	
NHML, large cell	14 (42%)	12 (44%)	
NKSCC, undifferentiated	17 (52%)	5 (18%)	
Other differential diagnosis	Germ cell tumor=1 (3%) Rabdomyosarcoma=1 (3%)	Neuroendocrine carcinoma=1(4%) Melanoma=1 (4%)	
Nodal			
Nasopharynx	6 (18%)	21 (78%)	
Tonsil and oropharynx	15 (45%)	-	
Cervical Lymph Node	3 (9%)	-	
Extranodal			
Nasal and paranasal spaces	5 (15%)	3 (11%)	
Maxilla and Mandible	2 (6%)	1 (4%)	
Buccal	1 (3%)	1 (4%)	
External Auditory Canal	1 (3%)	1 (4%)	

From the image acquisition, a total of 7,131 images were obtained, namely 3,893 images from large cell NHML and 3,238 images from nasopharyngeal carcinoma cases. The mean of images taken for each large cell NHML preparation was 118 images (standard deviation/SD: 104 images), whereas in the nasopharyngeal carcinoma group it was 120 images per preparation (SD: 155 images). The variation in the number of images is due to the variation in the tissue area of each case and the components relevant to the diagnosis. There was no significant difference in the number of images per preparation of the two groups (p=0.47). Each image is 1440x1080 pixels with each pixel representing an area of 0.21x0.21 micrometers. To prevent the dominance of cases with multiple slides in the model formulation, each case is limited to a maximum of 274 images. This figure is derived from the sum of the 3rd quartile of the distribution of

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

images per preparation with 1.5 times the IQR of the distribution of images per preparation.

After image segmentation, 3.1 million objects were identified in the large cell NHML group and 2.4 million objects in the nasopharyngeal carcinoma case group. After elimination of cells with a size of less than 10 pixels, 2.8 million objects and 2.2 million objects remained in large cell NHML and nasopharyngeal carcinoma, respectively. The mean reduction ratios for the large cell NHML and nasopharyngeal carcinoma groups were 8.86% (SD=2.07%) and 7.89% (SD=2.32%), respectively. After reduction, the mean of large cell NHML was 731 objects per image (SD=122 objects) while the mean of nasopharyngeal carcinoma was 686 objects per image (SD=134 objects). The difference in the number of objects before elimination, after elimination, and the reduction ratio between the two case groups was significant (p<0.01).

Table 3. Mean and standard deviation from most selected 20 factors with various value of alpha.

		NHML, large cell		NC*		
Factors	Unit	Mean	SD	Mean	SD	p
Minimum global intensity	%	32.29	2.62	28.73	3.93	0.00
3rd quartile global intensity	%	77.45	2.45	74.51	3.00	0.00
3rd quartile nuclear minimum diameter	µm	5.81	0.48	6.04	0.35	0.00
1st quartile nuclear solidity	%	71.37	2.57	67.95	3.67	0.00
3rd quartile nuclear circularity**	%	42.43	15.92	43.54	9.28	0.06
Median nuclear perimeter	µm	82.33	5.53	81.53	3.23	0.00
Minimum Cellular area	µm2	1.94	0.38	1.96	0.21	0.00
3rd quartile nuclear circularity	µm	21.95	2.13	22.41	1.40	0.00
Median nuclear circularity**	%	29.85	4.02	30.25	2.27	0.61
1st quartile global intensity	%	64.51	3.09	60.55	4.36	0.00
Median global intensity	%	15.55	58.50	19.52	6.70	0.00
Range global intensity	%	1.52	0.23	1.52	0.13	0.00
Minimum cellular maximum diameter	µm	5.01	2.40	5.95	1.31	0.00
Median nuclear solidity	%	88.85	4.15	88.24	2.54	0.00
Median nuclear minimal diameter	µm	4.55	0.36	4.81	0.33	0.00
1st quartile nuclear diameter ratio	%	1.35	0.06	1.32	0.04	0.00
Minimum cellular minimal diameter	µm	78.45	3.33	76.59	2.21	0.00
Median nuclear mathematical eccentricity		3.16	0.93	3.53	0.66	0.00
Median nuclear extent*	%	62.38	4.89	62.27	2.75	0.21
1st quartile nuclear extent*	%	54.56	5.31	54.53	2.89	0.71

*Nasopharyngeal Carcinoma
**Highlighted factors shows no significant difference between NHML and nasopharyngeal carcinoma (p>0.05)

There were 161 factors derived from the minimum value, 1st quartile, median, 3rd quartile, maximum value, range, and IQR of each object property as listed in table 1. Binary logistic regression per individual factor selects 76 factors with significant p>0.05 in differentiating the two groups of cases. The distribution of these factors is shown in Table 2. A combination of factors with a number of factors varying from 1 to all factors (without selection) was obtained by varying the alpha parameter from the lasso regression.

It can be seen that of the 20 factors that are most often selected from the lasso regression, it is found that the nuclear factor has a dominant role. Most of the factors are dominated by nuclear properties, namely as many as 12 out of 20 factors. A total of 5 factors are derived from the intact image properties. The remaining 3 factors come from cellular property factors. There were no factors of cytoplasmic properties in the 20 selected factors. As many as 5 of the 7 intensity factors are included in the 20 selected factors. Factors related to nuclear shape such as nuclear circularity, nuclear diameter ratio, nuclear eccentricity, and nuclear extension appear 6 times. However, 4 of them chose a non-significant difference (p<0.05). A total of 4 factors are related to the size of the nucleus, including the minimum diameter of the nucleus and the circumference of the nucleus. As many as 2 factors of nuclear solidity can assess the degree of nuclear vesicularity

The selected combination of factors is used to form a binary logistic regression model. The accuracy of the model is shown in Figure 2.

Changes in accuracy have a higher ompact when the number of factors from the model is less. This can be caused when the number of factors is small, only the most relevant factors are selected. Conversely, when the number of factors increases, other additional factors only play a role in correcting the factors that are more relevant.

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

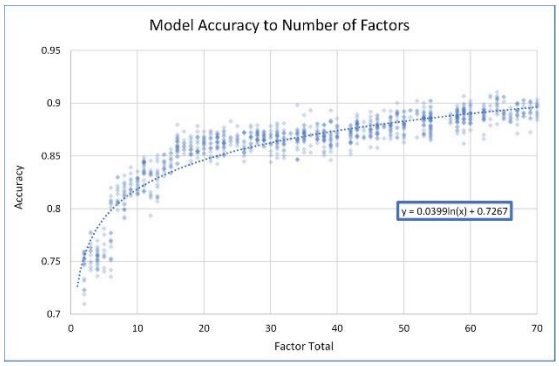

Figure 2. Indicates the level of accuracy of the model with various number of factor combinations. Each point represents one binary logistic regression model.

The sensitivity, specificity, and accuracy of the models with combinations of all factors, 50 factors, 20 factors, 10 factors, and 5 factors are shown in Table 4. The predictions of the various models in confirmed cases of large cell NHML (81.8%, SD=24 .4%) and confirmed cases of carcinoma (84.4%, SD = 13.9%). The sensitivity of the model is higher than the specificity but the difference between the two is not significant (p=0.30). Decreasing the value of the sensitivity and specificity of the model is positively correlated with the fewer factors used.

Table 4. Contingency tables, sensitivity, specificity, and accuracy of the model on a combination of all factors, 50 factors, 20 factors, 10 factors, and 5 factors.

		NHML	Carcinoma	Accuracy	Sensitivity	Spesificity	P Value
		QH all factors (76 factors)					
IHC	NHML	806	82	89.86%	91.38%	87.89%	>0.01
	Carcinoma	76	595				
		QH 50 selected factors					
IHC	NHML	791	97	87.88%	89.58%	85.65%	>0.01
	Carcinoma	92	579				
		QH 20 selected factors					
IHC	NHML	777	111	87.17%	89.72%	83.98%	>0.01
	Carcinoma	89	582				
		QH 10 selected factors					
IHC	NHML	762	126	83.32%	85.04%	81.00%	>0.01
	Carcinoma	134	537				
		QH 5 selected factors					
IHK	NHML	732	156	81.08%	84.04%	77.32%	>0.01
	Carsinoma	139	532				

QH = *Quantitative Histomorphometry*

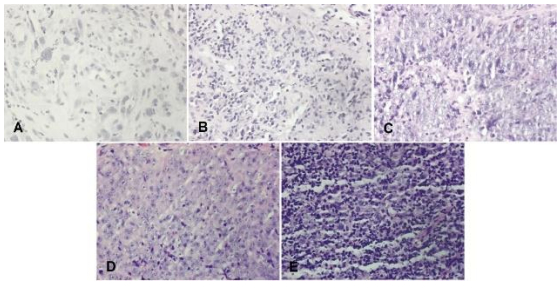

Figure 3. A. NHML predicted as a carcinoma. B. Carcinoma predicted to be NHML especially in multi-factor models. C. Carcinoma predicted to be large cell NHML especially in low-factor models. D. Typical nasopharyngeal carcinoma and E. Typical large cell NHML. The accuracy of all models on preparations D and E is more than 90%.

There are several cases showing different prediction patterns from other cases. In the NHML case, there were cases with much lower predictions with a magnitude of 9.5% (z-score=-2.96, p<0.01). This low prediction appears in all models with various combinations of factors. In this case, histopathological features of hypocellular tissue, edematous stroma, and prominent pleomorphism were seen (Figure 3 inset A).

In the case of carcinoma, poor accuracy was found in two cases of nasopharyngeal carcinoma. The first case is shown in inset B of figure 3 with an accuracy of 47.6% (p < 0.01). The addition of the factors involved in the model actually lowers the model's prediction accuracy from 55% on 5 factors, 46% on 10 factors, 50% on 50 factors, up to 33% on all selected factors.

The second case of nasopharyngeal carcinoma (shown in Figure 3 inset C) had an

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

accuracy of 52.4% (P=0.02). In this case, the increase in the factors involved in the more factor model, the accuracy increases from 24.2% in the 5 factor model, 47.1% in 10 factors, 58.1% in 20 factors, 64.0% in 50 factors up to 68.4% on all factors.

DISCUSSION

Diagnostic patterns of HE preparations

Based on the histopathological examination of hematoxylin and eosin (HE), the sensitivity and specificity of the primary diagnosis in large-cell NHML cases are 46% and 78%, respectively. The accuracy of the primary diagnosis of HE is 60%. Several studies showed similar results. A study from the Lymphopath Network found a general diagnosis accuracy of 80.3% for hematolymphoid and up to 62.2% in second opinion cases. The results of a study by Chang et al in Taiwan showed a misclassification rate of up to 60% (242 out of 406 cases) and an accuracy of 40%. In the study by Chang et al, the main cause of misclassification is ambiguous diagnosis.[14] A study from Cerrahpaşa Faculty of Medicine, Turkey showed an incidence of misclassification of up to 45.6% (774 of 1698 cases) and an accuracy of 54.4% of cases with misclassification of T-cell/ Histiocyte Rich B-cell Lymphoma to Hodgkin's lymphoma in 24 cases. In this study, 5 out of 33 cases of large cell NHML also had a primary diagnosis of Hodgkin's lymphoma.[15]

The sensitivity of HE diagnosis by looking at the differential diagnosis to identify NHML is 88% with a specificity of 96%. The accuracy of the HE diagnosis accompanied by a differential diagnosis is 92%. Most cases are accompanied by a secondary diagnosis. This demonstrates that both diagnoses in the study were considered by pathologists and confirms the limitations of conventional histopathological examination.[14] Only 9% of cases of large cell NHML are without a secondary diagnosis. In cases of nasopharyngeal carcinoma, the number of cases without a differential diagnosis is higher (30%). This indicates that the ability to identify nasopharyngeal carcinoma is better than identifying large cell NHML.

How quantitative histomorphometric analysis works in differentiating large cell NHML from nasopharyngeal carcinoma

Quantitative histomorphometric analysis works by measuring features that can only be assessed qualitatively by the human eye. The measurement results were then compared on the clinicopathological features studied. Several other studies have used histomorphometrical analysis to predict subvisual properties such as gene expression.

The study by Lu C et al in the Yale University pathology archive showed that the extraction results of 45 features of the shape and texture of the nucleus and 13 features of the orientation of the nucleus from HE preparations could predict 10-year survival in breast cancer with positive ER immunophenotype and not yet metastasized to regional lymph nodes. Lu C et al's study offers an alternative to molecular assays such as Oncotype DX, Mammaprint, and PAM50. Lu C et al's study obtained an accuracy of 61 – 70%. Lu C et al attributed increased heterogeneity of nuclear orientation and nuclear irregularity to the aggressiveness of the lesions. However, in the study of Lu C et al, no analysis was carried out for each feature individually.[16]

Whitney J et al's study predicts the molecular subtype of medulloblastoma by measuring features of chromatin texture, nucleus shape and the relation of the distance from one nucleus to another nucleus in 46 medulloblastoma patients. Features of chromatin texture and nuclear shape are used because they are related to anaplasia.[17] To date, no study has been found that utilizes quantitative histomorphometry to differentiate large cell NHML from undifferentiated carcinoma.

According to the article by Hamilon PW et al, digital image analysis such as quantitative histomorphometry works by identifying differences that are too small to be observed by the human eye..[18] The difference in global intensity between large cell NHML and nasopharyngeal carcinoma was only 3.6% ($p < 0.01$). Factors of global intensity alone were the best performing factors in differentiating these two groups of cases. The global intensity factor accuracy can reach more than 70% accuracy. This accuracy is higher than the primary diagnosis of HE.

Other examples of differences in factors between the two groups are the minimum diameter of the nucleus with a size of 0.23 micrometers, the nuclear solidity with a size of 3.4%, and the smallest cell area with a difference of 3.98 μm2. All of these differences were significant ($p<0.01$).

However, a factor that was not significantly different ($p>0.05$) was found in the 20 selected features in Table 3. This factor was the 3rd quartile of nuclear circularity ($p=0.06$) in the 5th position, median nuclear circularity ($p=0.61$) at the 9th position, median nuclear extension ($p=0.21$) and 1st quartile of nuclear

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

extension (p=0.71). The involvement of factors that do not differ significantly in the model indicates that each factor does not work with other factors independently. Each factor corrects the other factor. This resembles the pathologist's assessment approach in stratifying tumor groups. For example, like the Nottingham score on breast carcinoma[19], FNCLCC score in sarcoma[20], or in the stratification of neuroendocrine neoplasm differentiation.[21] In this scoring, each scoring factor does not stand alone and corrects the deficiencies of other factors.

The role of the pathologist in computational pathology and the limitations of quantitative histomorphometric analysis

Histomorphometric analysis is a measurement of tissue components. In daily practice, it is generally done qualitatively or semi-quantitatively. In computer-assisted quantitative histomorphometrical analysis, the results of the analysis are more detailed, fast, inexpensive, and can be replicated.[9,22] This tends to lead to the erroneous understanding that the pathologist no longer has a role in histopathological assessment.

Basically, quantitative histomorphometric analysis is a tool and each tool has a specific function. In this study, the analysis was limited to only two case groups. Utilization outside the case cluster will require a different model. This study also restricted the analysis to only selected histopathological images that were relevant to the final diagnosis. The determination of this relevance must be owned by the user of the model, namely the pathologist. Irrelevant images can lead to wrong conclusions. The narrow spectrum of the model must be well understood by the pathologist. This approach is referred to as human-in-the-loop. Boden et al's study shows that using the human-in-the-loop approach can correct errors in the algorithm such as determining segmentation in bad staining preparations and segmentation errors between tumor and stroma.[23]

The model is formulated by looking for the most common patterns and can represent as much data as possible into the algorithm. As a result, the model cannot work on morphological features that have a different pattern than most. In this case, which was found in three cases, it is still necessary to have an independent molecular examination of morphology. Quantitative histomorphometric examination supports HE examination as long as the preparation is of good quality. Quantitative histomorphometry is used as an initial screening before proceeding with the IHC examination. The IHC examination carried out can be adjusted based on the predicted results of the model so that the selected IHC can be more efficient.

CONCLUSION

Quantitative histomorphometry can increase the concordance between the histopathological diagnosis of HE and the results of IHC examination in differentiating large cell NHML from undifferentiated nasopharyngeal carcinoma. Quantitative histomorphometry on HE preparations has a higher accuracy than the histopathological examination of HE.

REFERENCES

1. Arber DA. Lymph Nodes. In: Goldblum JR, Lamps LW, McKenney JK, Myers JL, ed. *Rosai and Ackerman's Surgical Pathology*. Eleventh E. Elsevier Inc.; 2018:1530–1631.
2. Sung H, Ferlay J, Siegel RL, Layersanne M, Soerjomataram I, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209–249.
3. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. Longo DL, ed. *N Engl J Med*. 2021;384(9):842–858.
4. Yoo KH, Lee H, Suh C. Lymphoma epidemiology in Korea and the real clinical field including the Consortium for Improving Survival of Lymphoma (CISL) trial. *Int J Hematol*. 2018;107(4):395–404.
5. Lin F, Liu H. Immunohistochemistry in Undifferentiated Neoplasm/Tumor of Uncertain Origin. *Arch Pathol Lab Med*. 2014;138(12):1583–1610.
6. WHO Classification of Tumours Editorial Board. *Head and neck tumours*. 5th ed. International Agency for Research on Cancer; 2022.
7. Disanto MG, Ambrosio MR, Rocca BJ, Ibrahim HAH, Leoncini L, et al. Optimal minimal panels of immunohistochemistry for diagnosis of B-Cell lymphoma for application in countries with limited resources and for triaging cases before referral to specialist centers. *Am J Clin Pathol*. 2016;145(5):687–695.
8. Rekhtman N, Bishop JA. *Quick Reference Handbook for Surgical Pathologists*. Springer Berlin Heidelberg; 2011.
9. Madabhushi A, Lee G. Image analysis and

P-ISSN 0215-7284
e-ISSN 25279106
Accredited by KEMENRISTEKDIKTI/Sinta-3

machine learning in digital pathology: Challenges and opportunities. *Med Image Anal*. 2016;33:170–175.
10. Whitney J, Dollinger L, Tamrazi B, Hawes D, Couce M, et al. Quantitative Nuclear Histomorphometry Predicts Molecular Subtype and Clinical Outcome in Medulloblastomas: Preliminary Findings. *J Pathol Inform*. 2022;13:100090.
11. Van Der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD,s et al. Scikit-image: Image processing in python. *PeerJ*. 2014;2014(1):1–18.
12. Muthukrishnan R, Rohini R. LASSO: A feature selection technique in predictive modeling for machine learning. In: *2016 IEEE International Conference on Advances in Computer Applications (ICACA)*. ; 2016:18–20.
13. Laurent C, Baron M, Amara N, Haioun C, Dandoit M, et al. Impact of expert pathologic review of lymphoma diagnosis: Study of patients from the French Lymphopath network. *J Clin Oncol*. 2017;35(18):2008–2017.
14. Chang C, Huang SW, Su IJ, Chang KC. Hematopathologic discrepancies between referral and review diagnoses: A gap between general pathologists and hematopathologists. *Leuk Lymphoma*. 2014;55(5):1023–1030.
15. Özkaya N, Başsüllü N, Demiröz AS, Tüzüner N. Discrepancies in Lymphoma Diagnosis Over the Years: A 13-Year Experience in a Tertiary Center. *Turkish J Hematol*. 2017;34(1):81–88.
16. Lu C, Romo-Bucheli D, Wang X, Janowczyk A, Ganesan S, et al. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab Investig*. 2018;98(11):1438–1448.
17. Whitney J, Corredor G, Janowczyk A, Ganesan S, Doyle S, et al. Quantitative nuclear histomorphometry predicts oncotype DX risk categories for early stage ER+ breast cancer. *BMC Cancer*. 2018;18(1):1–15.
18. Hamilton PW, Bankhead P, Wang Y, Hutchinson R, Kieran D, et al. Digital pathology and image analysis in tissue biomarker research. *Methods*. 2014;70(1):59–73.
19. Zhou L, Rueda M, Alkhateeb A. Identifying biomarkers of nottingham prognosis index in breast cancer survivability. In: *Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM; 2021:1–9.
20. Schwen LO, Nitsch J, Bauer S, Bertram S, Goetz M, et al. 1659P AI-based grading approach identifies FNCLCC grade 3 soft tissue sarcomas. *Ann Oncol*. 2020;31:S988.
21. Kyriakopoulos G, Mavroeidi V, Chatzellis E, Kaltsas GA, Alexandraki KI. Histopathological, immunohistochemical, genetic and molecular markers of neuroendocrine neoplasms. *Ann Transl Med*. 2018;6(12):252–252.
22. Pantanowitz L, Sharma A, Carter AB, Kurc T, Sussman A, et al. Twenty years of digital pathology: An overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. *J Pathol Inform*. 2018;9(1).
23. Bodén ACS, Molin J, Garvin S, West RA, Lundström C, et al. The human-in-the-loop: an evaluation of pathologists' interaction with artificial intelligence in clinical practice. *Histopathology*. 2021;79(2):210–218.

Maj Patol Indones 2024; 33(1): 740-748
DOI: 10.55816/mpi.v33i1.526