Semi-Automatic Rating Method for Neutrophil Alkaline Phosphatase Activity

Kanae Sugano,^{1*} Kotomi Hashi,^{2*} Misaki Goto,^{3*} Kiyotaka Nishi,^{4*} Rie Maeda,^{5*} Keigo Kono,⁶ Mai Yamamoto,⁶ Kazunori Okada,⁷ Sanae Kaga,⁷ Keiko Miwa,^{7,8} Taisei Mikami,⁷ and Nobuo Masauzi^{7†}

¹Department of Clinical Laboratory, NTT East Japan Sapporo Hospital, Sapporo, Japan

²Department of Clinical Laboratory, Daiichi Kishimoto Medical Laboratory Center, Sapporo, Japan

³Department of Clinical Laboratory, Sapporo Hokuyu Hospital, Sapporo, Japan

⁴Department of Clinical Laboratory, Teine Keijinkai Hospital, Sapporo, Japan

⁵Department of Clinical Laboratory, KKR Sapporo Medical Center, Sapporo, Japan

⁶Graduate school of Health Sciences, Hokkaido University, Sapporo, Japan

⁷Department of Medical laboratory Science, Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

⁸Department of Cardiovascular Regenerative Medicine, Osaka University Graduate School of Medicine, Suita, Japan

Background: The neutrophil alkaline phosphatase (NAP) score is a valuable test for the diagnosis of myeloproliferative neoplasms, but it has still manually rated. Therefore, we developed a semi-automatic rating method using Photoshop® and Image-J, called NAP-PS-IJ. **Methods:** Neutrophil alkaline phosphatase staining was conducted with Tomonaga's method to films of peripheral blood taken from three healthy volunteers. At least 30 neutrophils with NAP scores from 0 to 5+ were observed and taken their images. From which the outer part of neutrophil was removed away with Image-J. These were binarized with two different procedures (P1 and P2) using Photoshop®. NAP-positive area (NAP-PA) and granule (NAP-PGC)

were measured and counted with Image-J. **Results:** The NAP-PA in images binarized with P1 significantly ($P < 0.05$) differed between images with NAP scores from 0 to 3+ (group 1) and those from 4+ to 5+ (group 2). The original images in group 1 were binarized with P2. NAP-PGC of them significantly ($P < 0.05$) differed among all four NAP score groups. The mean NAP-PGC with NAP-PS-IJ indicated a good correlation ($r = 0.92$, $P < 0.001$) to results by human examiners. **Conclusions:** The sensitivity and specificity of NAP-PS-IJ were 60% and 92%, which might be considered as a prototypic method for the full-automatic rating NAP score. J. Clin. Lab. Anal. 31:e22022, 2017. © 2016 Wiley Periodicals, Inc.

Key words: granule counting; digital image processing; Tomonaga's method; Image-J; Photoshop

INTRODUCTION

The neutrophil alkaline phosphatase (NAP) is an enzyme in cytoplasmic granules of neutrophil. Kaplow's (1) and Tomonaga's (2) methods are recommended by the International Counsel for Standardization of Hematology as cyto-chemical stain methods for NAP (3). The NAP activities in a neutrophil are rated based on the number and distribution of NAP-positive granules in the cytoplasm. The total NAP activity score (NAPScore) is calculated as the sum of the activity scores in each cell (NAPScoreCell) and their cell

number. The NAPScoreCell is based on five levels (from 0 to 4+) with Kaplow's method, which is used mainly in western countries, or six levels (from 0 to

*Equally contributed to this study.

†Correspondence to: Nobuo Masauzi, Department of Medical laboratory Science, Faculty of Health Sciences, Hokkaido University, North 12 – West 5, Kita-ku, Sapporo, 060-0812, Japan. E-mail: nmasauzi@hs.hokudai.ac.jp

Received 9 May 2016; Accepted 4 June 2016

DOI 10.1002/jcla.22022

Published online in Wiley Online Library (wileyonlinelibrary.com).

5+) with Tomonaga's method, which is used mainly in Japan (3). The NAPScore is valuable information for distinguishing myeloproliferating neoplasms (MPNs), especially between chronic myelocytic leukemia and polycythemia Vera (PV), and has been used in the diagnosis of paroxysmal nocturnal hematuria (4). Recently, as the cytogenetic or gene test becomes more practical, these examinations for MPNs are routinely used. Consequently, the NAPScore in the diagnosis of MPNs is now supplementary. However, a significant decrease in the NAPScore has been reported in PV with JAK2V617F mutation (5). Therefore, the NAPScore is still valuable due to reduced cost and lighter equipment for the examination, especially in developing countries. The NAPScoreCell is still determined by the naked eyes in clinical laboratories. Therefore, objectivity and reproducibility are low, and the result is dependent on the subjective judgment of each examiner. Automatically determining the NAPScoreCell provides more stable results. Such an automatic method is also expected to be applied for evaluating or counting other positive granules such as iron stain-positive granules in myelodysplastic syndrome (MDS) or sideroblastic anemia. Therefore, we developed a semi-automatic image analysis method with widely used image processing software, which is a prototype for fully automatic analysis with special software.

MATERIALS AND METHODS

Neutrophil Alkaline Phosphatase Stain and Image Files of Neutrophil

Peripheral blood was collected from three healthy volunteers (mean age: 21.3 years old, three females) after obtaining informed consent. Two blood films from each donor for NAP staining were smeared with the wedged method on silicon-courted slide glass (#1-3346-01; AS ONE Corporation, Osaka, Japan). The NAP staining was done with Tomonaga's method (2) using an alkaline phosphatase staining kit ('ARU-HOSU' staining kit[®]; MUTO Chemical Industries, Ltd., Tokyo, Japan) following the staining instructions provided by the manufacturer. From one blood film from each donor, more than 30 neutrophil of each group tentatively judged as NAPScoreCell from 0 to 5+ (3) were taken their images with a digital camera for microscopes (Dino-Eye[®], 1.35 M pixels; ANMO Co. Inc., New Taipei City, Taiwan). A total of over 180 JPEG image files were stored in a folder called the standard image folder (SIF). The other blood films were observed according to routine NAP activity rating, i.e., the sum of NAPScoreCells of 100 neutrophils in one blood film. A total of 100 neutrophil images

from each blood film were continuously taken in the order of appearance. These images were stored as another set of JPEG files in a folder called the objective image folder (OIF). The NAP rating for every neutrophil image in the SIF was evaluated again by five senior students in our laboratory according to the usual criteria for the rating (3). Every neutrophil image in the SIF was finally rated based on its NAPScoreCell from 0 to 5+ determined by the average integer of six values rounded off to the first decimal place.

Image Processing Method

Photoshop[®] CS6 (Adobe System Inc., San Jose, CA) (PS) and Image-J (Ver.1.43, available at rsb-web.nih.gov/ij/; accessed April 14, 2014) (IJ), hereafter called the "PS-IJ method", were used for image processing. The details of the original "PS-IJ method" are described elsewhere (6). We developed a modified version of the PS-IJ method for NAP rating called the "NAP-PS-IJ". The outer part and nucleus in each neutrophil image were removed using the "Edit > ClearOutside" and "Edit > Clear" tools in IJ and these areas were filled with the brightest intensity. Before this procedure, the margins of both areas were determined by the naked eye and the lines were drawn by hand using a mouse or pen tablet (Bamboo[®]; WACOM Co. Inc., Tokyo, Japan). These images were processed with two newly developed binarization procedures (P1 and P2) (Fig. 1) using the "Image > Adjustment > Tone Curve" tool in PS. The P1 procedure involves binarizing the original image to an optimum one to distinguish the NAP-positive area from the entire cytoplasm in the neutrophil. The areas of NAP-positive parts and entire cytoplasm were relatively measured with the number of pixels using the "Analyze > Histogram" tool in IJ. This tool outputs a file including a list of pixels based on relative intensity. The P2 procedure involves binarizing the original image to an optimum one with which it is easiest to count the NAP-positive granules in the cytoplasm of the neutrophil. The NAP-positive granules in the cytoplasm were counted using the "Analyze > AnalyzeParticles" tool in IJ.

Threshold Value for Distinguishing Neutrophil Alkaline Phosphatase Rating

For determining the threshold values for automatically obtaining the NAPScoreCell, the ratio of the NAP-positive area to the entire cytoplasm (positive area ratios: PAR) was calculated and the number of NAP-positive granules (granule count: GC) was counted for every image in the SIF. Next, the averages

Fig. 1. Two newly developed binarization procedures (P1 and P2). Horizontal axes indicate relative intensity (RI) of pixels in image before binarization. Vertical axes of graphs A and C indicate RI after binarization with P1 or P2, respectively. Line in Graph B indicates original linear relation of RI between original and processed images. Lines in Graphs A and C indicate binarization with P1 or P2, respectively. RI of pixel changed to the brightest RI (= 255) or darkest RI (= 0) with procedure P1 or P2, when RI of the pixel was brighter or darker than each threshold RI (TV), respectively. In this figure, brightest areas indicated in red and darkest areas indicated in black. The center of three cell images was stained using neutrophil alkaline phosphatase stain at the 1,000 $\times$ magnification. TV, threshold value; RI, relative intensity,

of these two values were calculated for every image in six groups of neutrophils with NAPScoreCells from 0 to 5+. The differences in the two averages of the PARs and GCs among the six groups were determined using the Kruskal–Wallis test (K-W test). The threshold values between two groups were determined using the test results.

Evaluation of NAPscore Cell and NAPScore Using NAP-PS-IJ

For images in the same NAPscore-cell group from 0 to 5+ in the SIF, the rate of agreement between the two methods (sensitivity) and specificity was first calculated, comparing the NAPScoreCells determined using the NAP-PS-IJ with those using the conventional method involving the naked eye. For three sets of 100 neutrophil images in the OIF, the NAPScores were then calculated using the conventional and NAP-PS-IJ and both values were compared. The numbers of NAP-positive granules from both methods were compared and the correlation between these results was evaluated.

Statistical Analysis

The differences in the results of PARs or GCs were determined using the Kruskal–Wallis (K-W) one-way analysis of variance (one-way ANOVA) test and Tukey's multiple comparison test. Spearman's correlation coefficient was used for evaluating the correlation

of GC with NAP-PS-IJ to those with conventional method. Statistical tests were calculated with EZR (7) and JMP[®] Pro 11 (SAS Institute Japan Ltd., Tokyo, Japan). All two-sided *P* value less than 5% is judged as significant.

RESULTS

The P1 binarizing procedure was applied for every neutrophil image in the SIF, and the PARs were calculated using the NAP-PS-IJ. The PARs (Fig. 2A) significantly ($P < 0.05$: K-W one-way ANOVA) differed from those of the images between two NAPScoreCell groups with scores from 0 to 3+ and from 4+ to 5+. However, the PARs (Fig. 2B) did not differ (NS: K-W one-way ANOVA) among the other pairs. The threshold value of the PAR for dividing groups between NAPScoreCells from 0 to 3+ and from 4+ to 5+ was determined by the median value (77.5%) between the 75th percentile of the former and 25th percentile of the latter.

Next, the P2 binarizing procedure was applied for the original pre-binarizing images, which had already been distinguished as NAPScoreCells from 0 to 3+ by the PAR of the P1 binarized images. For the resulting images, the NAP-positive granules in the cytoplasm were counted using the “AnalyzeParticle” (AP) tool in IJ. The GCs (Fig. 2C) significantly ($P < 0.05$: K-W one-way ANOVA) differed among all four NAPScoreCell groups. Those of each pair (Fig. 2C) also significantly ($P < 0.005$ or $P < 0.0001$: Tukey's

At this point in this study, neither the PAR significantly (NS: K-W one-way ANOVA) differed between

The NAP-positive GCs in the cytoplasm of 328 images, with which the NAPScoreCells were from 0 to 2+, were determined using the conventional method. These significantly ($r = 0.82$, $P < 0.001$) correlated (Fig. 4A) with those from the conventional method.

Fig. 4. The evaluation results of NAP-PS-IJ. (A) Correlation of NAP-positive granule count using NAP-PS-IJ with that using conventional method is illustrated. Small dots indicate number of NAP-positive granules counted using NAP-PS-IJ (the vertical axis) and conventional method (horizontal axis) for 328 images with NAP scores from 0 to 2 by using conventional method. A line in graph indicates regression line for both GCs. (B) Comparison of NAP scores determined with conventional method and those with NAP-PS-IJ is illustrated. Circles indicate score determined by examiner. Bars indicate average scores from five examiners. Triangles indicate score from NAP-PS-IJ method. NAP, Neutrophil alkaline phosphatase; GC, Granule count; NAP-PS-IJ, the method for NAP rating using Photoshop® and Image-J.

TABLE 1. Distinction Results and the Performance Values of NAP-PS-IJ

	Conventional human-eye method						Total	Sensitivity	Specificity
	Score								
	0	1	2	3	4	5			
NAP-PS-IJ method									
Score									
0	78	32	1	0	0	0	111	0.690	0.947
1	24	59	16	1	0	0	100	0.573	0.935
2	2	11	71	19	2	1	106	0.634	0.944
3	0	0	15	97	8	0	120	0.542	0.958
4	5	1	7	51	52	33	149	0.464	0.844
5	4	0	2	11	50	80	147	0.702	0.892
Total	113	103	112	179	112	114	733		

NAP-PS-IJ, the method for NAP rating using Photoshop® and Image-J.

The low indicates the numbers of images in each NAP score of cell decided with NAP-PS-IJ. The column indicates those with conventional human-eye method. The column of the sensitivity or specificity indicates the rates for images in every NAP score of cell decided with NAP-PS-IJ, comparing with the result with conventional human-eye method. The average of sensitivity and specificity is 0.60 and 0.92, respectively.

For the neutrophil images in the SIF, the number of images with every NAPScoreCell determined using the PS-IJ method is listed in Table 1 for every image with that determined with the conventional method. The sensitivity and specificity of the NAPScoreCells determined using the NAP-PS-IJ with threshold values determined in this study (Table 1) were calculated, and the average of those are 0.60 and 0.92, respectively.

For three sets of 100 neutrophil images in the OIF, the NAPScores were calculated using the conventional and NAP-PS-IJ, and both values were compared. On every slide, the NAPScores with the NAP-PS-IJ were 6–17% higher than those with the conventional method (Fig. 4B). And the time required for calculated NAPScore of 100 neutrophil images using NAP-PA-IJ

was almost 155 min, which was comprised of the following six sub-processes, i.e., 15 min for taking pictures, 50 min for determining the outer margins of the cytoplasm and nacreous, 25 and 20 min for binarizing images with procedure P1 and P2, respectively, 20 min for counting NAP-positive granules with IJ, and 25 min for measuring area of both NAP positive and cytoplasm of each cell and calculating NAP-positive ratio.

DISCUSSION

We developed a semi-automatic rating method for NAP activity, i.e., the NAP-PS-IJ, which indicates similar values and correlation with the conventional

method involving the naked eye. By using the NAPScoreCell determined with the NAP-PS-IJ, the NAP-Scores were quite similar to those determined with the conventional method, indicating no significant difference (6–17%). For using NAP-PS-IJ, no special proficiency was needed because Photoshop® is one of the most popular image processing software, and Image-J is also routinely used in laboratory all over the world. As indeed this presented study had been performed mainly by students in undergraduate class.

To the best of our knowledge, there have been only two reports concerning the automatic rating of NAP-Scores (8, 9). There were no concrete descriptions of a method for NAP rating, in the former, of which the value and range of NAPScores were not similar to those with the conventional method, indicating a reverse correlation. In the near future, these automatic image analyzing tools will be improved and be commonly applied to microscopic observation systems. In such situations, it is important and more convenient that a conventional human-eye rating method and the results are followed suit by such newly developed automatic tools. Therefore, the NAP-PS-IJ is more practical than the previously reported methods.

However, the NAP-PS-IJ exhibited lower sensitivity in all scores for practical use. The sensitivities in NAPScoreCell of 3+ and 4+ were especially lower than those of the other scores with the NAP-PS-IJ. As shown in Table 1, the numbers of images with NAPScoreCells of 3+ or 4+ were 179 or 112 and 120 or 149 with the conventional and NAP-PS-IJ, respectively, indicating a significant difference between the two NAPScoreCell groups. Among 149 cells that were judged with NAPScoreCells of 4+ with the NAP-PS-IJ, only 52 (35%) were correctly determined as the same NAPScoreCell with the conventional method, but almost the same number (51) was incorrectly rated with the NAP-PS-IJ, which were rated with a NAPScoreCell of 3+ with the conventional method. When P1 divided neutrophil images into two groups, i.e., a group with NAPScoreCells from 0 to 3+ and those with 4+ and 5+, many images in the NAPScoreCell 3+ group were distinguished into the other cell groups, in which the NAPScoreCells were 4+ and 5+. This is why the sensitivity in cells with NAPScoreCells of 3+ and 4+ are relatively lower than in those with other NAP ratings.

These results indicate that the distinction between neutrophil images with NAPScoreCells of 4+ and 5+ is impossible by measuring the NAP-positive area in the cytoplasm. Therefore, the cell images with NAPScoreCells of 4+ and 5+ were divided based on tentative

threshold without any ground, i.e., the median value of both average values of the PAR. These might be limitations of the NAP-PS-IJ, which uses the PAR and GC as tools for rating NAPScoreCell. A more complicated method should be used for automatic rating of NAP activities, such as texture image analysis methods.

We had tried to apply fast Fourier transformation for neutrophil images with NAPScoreCells of 4+ and 5+, which unsuccessfully resulted in no statistically significant difference. A different texture analysis method is needed to distinguish between such images.

AUTHOR CONTRIBUTIONS

NM is the principal investigator and takes primary responsibility for this study. KS, KH, MG, KN, and RM equally conducted all laboratory work for this study. KS statistically analyzed data and made all graphs and tables. KK, MY, KO, SK, KM, and TM attended every research meeting and gave many valuable comments for both this study and paper. KS made a draft for this manuscript in Japanese, and NM translated the draft into English.

REFERENCES

1. Kaplow LS. Leukocyte alkaline phosphatase cytochemistry: Applications and methods. *Ann NY Acad Sci* 1968;155:911–947.
2. Tomonaga M, Sakai T, Okuzaki M. Studies on leukocyte alkaline phosphatase, 1. Use of Naphthol AS-Mx phosphate-Fast Blue RR staining method. *Acta Haemat Jap* 1963;26:179–192.
3. Shibata A, Bennet JM, Castoldi GL, et al. Recommended methods for cytological procedures in haematology. *Clin Lab Haemat* 1955;7:55–74.
4. Bain BJ. 2015. Important supplementary tests. In Bain BJ, (ed.). *Blood Cells a Practical Guide*. West Sussex, UK: John Wiley & Sons Ltd. p 277–294.
5. Okabe M, Yamaguchi H, Usuki K, et al. Clinical features of Japanese polycythemia vera and essential thrombocythemia patients harboring CALR, JAK2V617F, JAK2Ex12del, and MPLW515L/K mutations. *Leukemia Res*, 2016;40:68–76.
6. Mizukami M, Yamada M, Fukui S, et al. Semi-automated segmentation and measurement of cytoplasmic vacuoles in a neutrophil with general-purpose image analysis software. *J Clin Lab Anal*, 2016. doi: 10.1002/jcla. 21957. [E-Pub ahead of print].
7. Kanda Y. Investigation of the freely available easy-to-use software ‘EZR’ for medical statistics. *Bone Marrow Transplant* 2013;48:452–458.
8. Tafas T, Cuckie HS, Nasr S, Krivchenia EL, Resvani E, Evans MI. An automated image analysis method for the measurement of neutrophil alkaline phosphatase in the prenatal screening of Down syndrome. *Fetal Diagn Ther* 1996;11:254–259.
9. Kaplow LS, Crouch JY, Hans Kunz JAM, Garcia GL. Assessment of leukocyte alkaline phosphatase by image analysis. *Ann NY Acad Sci* 1986;468:85–92.