Application of Image Analysis and Neural Networks to the Pathology Diagnosis of Intraductal Proliferative Lesions of the Breast

Noriyoshi Fukushima,[1] Hiroyuki Shinbata,[2] Takahiro Hasebe,[1] Tomoyuki Yokose,[1] Atsushi Sato[2] and Kiyoshi Mukai[1, 3]

[1]Pathology Division, National Cancer Center Research Institute, East, 6-5-1 Kashiwanoha, Kashiwa 277 and [2]Pattern Recognition Research Laboratory, Information Technology Research Laboratories, Nippon Electric Company Corporation, 4-1-1 Miyazaki, Miyamae-ku, Kawasaki 216

We studied whether a computer-assisted system using a combination of data collection by image analysis and analysis by neural networks can differentiate benign and malignant breast lesions. Forty-six intraductal lesions of the breast were studied by pathologists and by the computer-assisted system. Histological evaluation was performed independently by three pathologists, and the lesions were classified into pathologically malignant ($n = 12$), undetermined ($n = 13$), and benign ($n = 21$). Computerized nuclear image analysis was performed using the CAS200 (Cell Analysis Systems, Elmhurst, IL) system to obtain data on nuclear morphometric and textural features. A neural network was constructed using the morphometric and texture data obtained from teaching cases of malignant and benign lesions. Then data for unknown cases were classified by the constructed neural network into neural network-malignant ($n = 11$), -undetermined ($n = 5$), and -benign ($n = 30$). The agreement rate between the diagnosis by pathologists and judgement by the computer-assisted system was 75%, excluding pathologically undetermined lesions. There were four false-negative but no false-positive results. False-negative cases had nuclei that were quite different from those of the teaching cases. The agreement rate obtained using either morphometric data or texture data only was lower than that using a combination of both. Selection of appropriate teaching data and incorporation of both morphometric and textural parameters seemed important for obtaining more accurate results. The present data suggest that development of a computer-assisted histopathological diagnosis system for practical use may be possible.

Key words: Pathology diagnosis — Image analysis — Neural network — Breast intraductal lesion

Histopathological diagnosis of intraductal proliferative lesions of the breast is often difficult. Differential diagnosis between epitheliosis and carcinoma *in situ*, and between papilloma and non-invasive papillary carcinoma, is especially troublesome.[1] Different pathologists may not always agree as to the malignant potential of a given lesion. Furthermore, quality control and standardization are difficult to achieve because histopathological diagnosis relies on the knowledge and personal experience of individual pathologists who evaluate the architectural and cytological features using light microscopy. Computer-assisted diagnostic decision support systems have recently been developed in certain medical fields,[2, 3] and neural networks have been considered a useful tool for quality control in diagnostic cytology[4-6] and pathology.[7-12] Neural networks computationally simulate the neurological processing of data by biological organisms.[7, 13] They characteristically have the ability to "learn" to associate a correct output classification with specific patterns of input data by training. The "knowledge" of a trained neural network is embodied in the strength of interconnections formed between processing elements, an equivalent of neurons in the brain. Possible applications of neural networks include pattern recognition, database search and knowledge extraction and decision-making. The pattern recognition ability has been applied to individual cell classification using cytometric quantities, such as size, shape, and texture.[7] The nuclear grading of breast carcinoma,[8] differential diagnosis of well differentiated hepatocellular carcinoma and benign lesions in cirrhotic liver[12] and evaluation of cervical smears[4] are some examples of neural network application to surgical pathology.

Computer-assisted diagnostic systems seem to excel in objectivity and reproducibility of evaluation. This advantage should be put into practical use for quality control and standardization of pathology diagnosis. In this report, we describe our attempt to examine the applicability of image analysis and neural networks to the diagnosis of intraductal lesions of the breast.

MATERIALS AND METHODS

Forty-six intraductal lesions of the breast were retrieved from the files of the Pathology Laboratory, National Cancer Center Hospital East. The original diag-

---

[3] To whom correspondence should be addressed.

noses of the selected lesions included epitheliosis ($n=5$), intraductal papilloma ($n=13$), atypical ductal hyperplasia ($n=11$), non-invasive ductal carcinoma ($n=5$), intraductal component of invasive carcinoma ($n=5$) and borderline papillary lesion ($n=7$). Comedo carcinoma was excluded because differential diagnosis of this lesion is not a problem. Formalin-fixed, paraffin-embedded tissue specimens were cut into 3-μm-thick sections for hematoxylin and eosin (HE) staining and 5-μm-thick sections for Feulgen staining with a "CAS" quantitative DNA staining kit (Cell Analysis Systems, Elmhurst, IL).[14] Selected lesions, ranging from 2 mm to 15 mm in diameter, were marked in the same area on both slides.

Histological evaluation was performed by three pathologists, each with more than 10 years of experience. Each lesion was evaluated according to not only nuclear find-

ings, but also structural features. The pathologists were requested to classify the lesions into three categories: benign, borderline, or malignant. Based on the results of their evaluation, the 46 ductal lesions were classified into three groups: pathologically malignant (PM), pathologically undetermined (PU), and pathologically benign (PB). There was no disagreement among the pathologists with regard to the PM and PB groups. The PU group included those lesions diagnosed as borderline and lesions for which a consensus diagnosis was not obtained.

A flow chart showing the path from image analysis to cell classification is shown in Fig. 1. Computerized nuclear morphological analysis was performed with a CAS200 image cytometer (Cell Analysis Systems) run by the cell measurement program. Scene segmentation was semi-automated. Slightly overlapping cells were manually divided by a boundary line using the cut-mode of the cell measurement program. We analyzed about 200 cells with an intact nucleus in each lesion. Twenty-five parameters including nuclear size, shape and density, and 22 Markov texture parameters were used to analyze the nuclear characteristics. Markovian texture analysis is a method of measuring optical texture based on gray-level transition probabilities in digitized images.[15] Each textural parameter was calculated for step sizes of 0.125 μm, and the digitized images had 256 gray levels.

We used data for one typical case each from the PM and PB groups as teaching data for the following experiments. In the malignant teaching lesion, papillotubular structures formed a cribriform pattern. The tumor cells had small to medium-sized round nuclei with a relatively fine chromatin pattern (Fig. 2A). This was a case of well differentiated or low-grade intraductal carcinoma. On

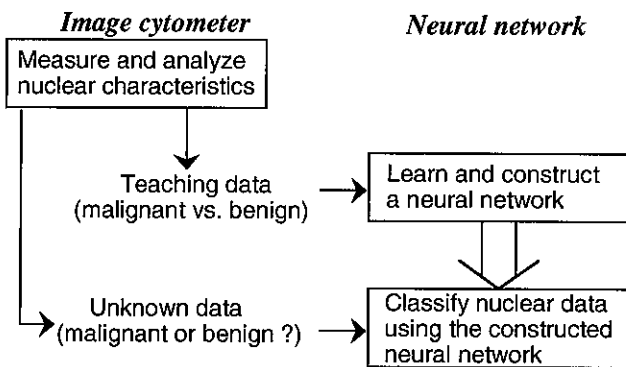*Image cytometer*   *Neural network*

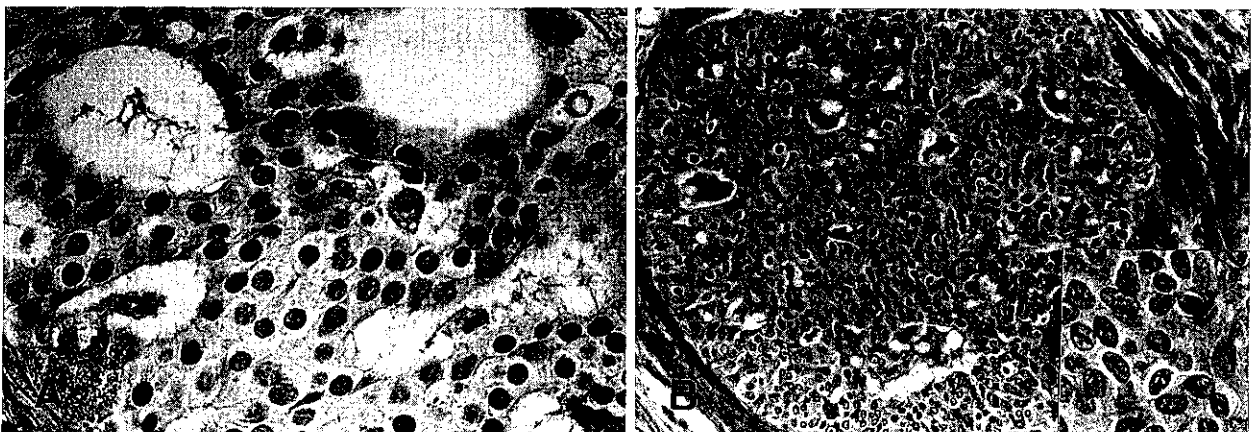Fig. 1. Flow chart of the examination.

Fig. 2. Histological findings of teaching cases. A, Malignant case shows a cribriform pattern. Nuclei are round to oval and the chromatin pattern is relatively fine. ×305. B, Benign case shows high cellularity and two cell patterns. Nuclei are oval and not arranged regularly. ×76 (inset ×305).

the other hand, the benign lesion, diagnosed as epitheliosis, showed solid and tubular growth with a fibrovascular stroma, as undefined cytoplasmic border, and oval nuclei with a relatively coarse chromatin pattern and/or grooving (Fig. 2B).

Neural network software (Neural Works, Professional II/plus, Neural Ware, Pittsburgh, PA) was instructed to construct neural networks using a three-layered back-propagation algorithm (malignant vs. benign). Several useful parameters for cell classification were selected by multiple regression analysis before feeding into the neural network. The number of input layer nodes was varied between 2 and 8, which was the number of selected parameters for the three kinds of cell-classification tests. The most accurate network for cell classification was obtained with 20 hidden nodes in the intermediate layer and 1,000 training times. The output layer consisted of two nodes representing diagnostic outcome, 0.0 for benign and 1.0 for malignant. According to the cell classification rate by the neural networks, the cases were divided into 3 groups: neural network-malignant (NM), -undetermined (NU), and -benign (NB). Lesions in which more than 60% of the cells were judged to be malignant by the neural networks were classified into the NM group, and those with less than 40% malignant cells into the NB group.

Cell classification tests were performed using three neural networks constructed with the same teaching sets
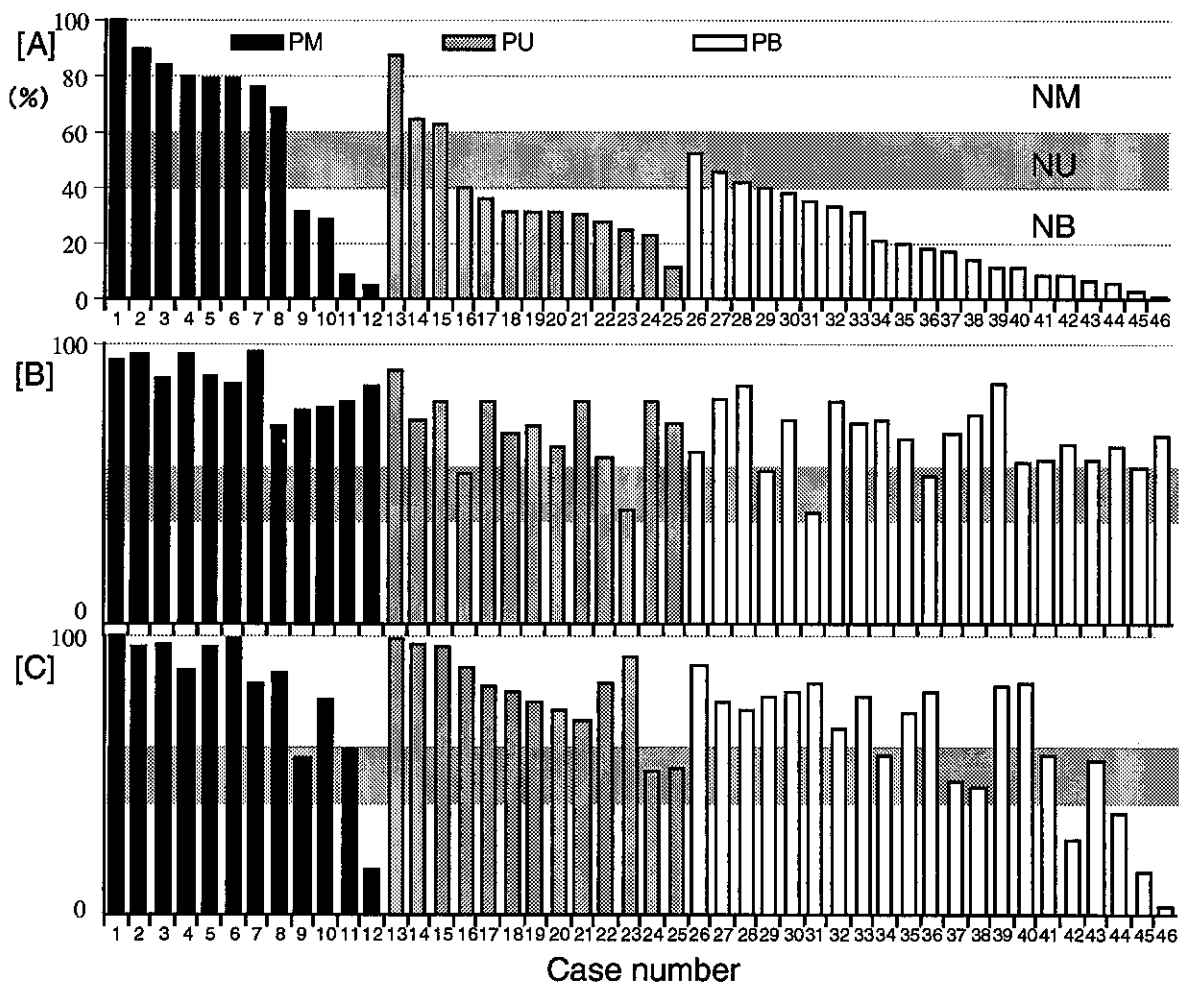

Fig. 3. Histogram of the frequency of "malignant cells" obtained using the neural network in 46 breast lesions. PM, pathologically malignant; PU, pathologically undetermined; PB, pathologically benign; NM, neural network-malignant; NU, neural network-undetermined; NB, neural network-benign. A, The neural network was constructed using combined morphometric and textural parameters; B, using morphometric parameters only; C, using five textural parameters.

and different parameters. In test A, the neural network was constructed using eight selected parameters including size, shape, and sum optical density, and five textural parameters (sum average, sum variance, difference entropy, information measure B, and coefficient of variation) (neural network A). In test B, only the morphometric parameters (size and shape) were used (neural network B), while in test C, only five textural parameters were used (neural network C).

RESULTS

There were 12 lesions in the PM group, 13 in the PU group, and 21 in the PB group (Fig. 3). All PM lesions except two were diagnosed as malignant in the original pathology reports; two lesions were called atypical ductal hyperplasia and borderline papillary lesion, respectively. The PU group included seven cases of atypical ductal hyperplasia, four intraductal papillomas, and two borderline lesions, while the PB group included three cases of atypical ductal hyperplasia, nine intraductal papillomas, six cases of epitheliosis, and three borderline lesions in the original pathology reports.

The results of analysis by neural network A showed that 11 lesions were classified into the NM group, five into the NU group, and 30 into the NB group (Fig. 3A). The rate of agreement between PM and NM was 66.7%, and that between PB and NB was 81.0% in test A. Four-fifths of the PU group were classified into the NB group. There were four false-negative (PM/NB) but no false-positive (PB/NM) results using the classification criteria of the present study (Fig. 3A). Histologically, nuclear chromatin tended to be more coarse in the false-negative
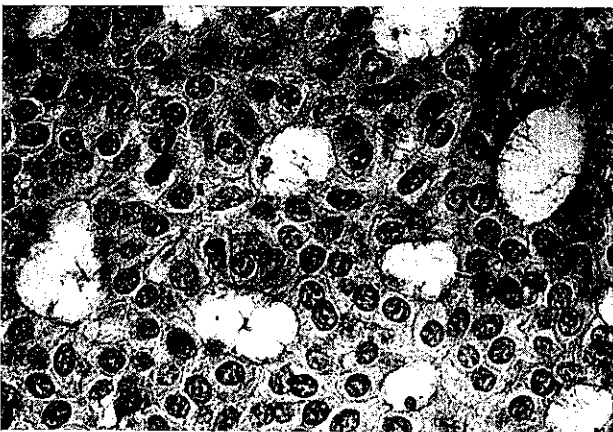

Fig. 4. Histological features of a "false-negative" case. Hyperchromasia and a coarse chromatin pattern are more notable than in a malignant teaching case. ×305.

lesions than in the PM/NM group (Fig. 4). Three PU lesions were classified into the NM group. In test B, there were 36 cases in the NM group, nine in the NU, and one in NB, and in test C there were 32 cases in NM group, nine in the NU, and five in NB (Fig. 3, B and C). There were many false-positive cases (PB/NP) in tests B and C; the rate of agreement between PM and NM was higher, but that between PB and NB was much lower than in test A.

DISCUSSION

Computerized diagnostic systems for pathology images require several useful parameters reflecting the histopathological features, i.e., nuclear and structural characteristics. In this study, we focused on analysis of nuclear features. Nuclear size, shape, density and chromatin pattern are important diagnostic clues used by pathologists for differential diagnosis between benign and malignant lesions of the breast.[1] These features may be represented in the morphometric and textural parameters to a certain degree.

The rate of agreement between pathological evaluation and classification by the computer-assisted system was about 75%, excluding borderline lesions (PU group). Histological variation in carcinoma seems to have been the major cause of the false-negative (PM/NB group) results. Carcinoma cells with a higher grade of atypia were not identified as malignant cells by neural network A, which was constructed using teaching data obtained from a case of well differentiated ductal carcinoma with a low grade of atypia and marked proliferative epitheliosis. The neural network was able to classify accurately cells which had nuclear characteristics similar to those of the teaching sets, but could not consistently classify cells which had characteristics different from the teaching cases. Dawson et al. demonstrated the feasibility of using image analysis as an objective means for nuclear grading of breast carcinoma.[8] They designed a neural network to classify the nuclear grade of breast carcinomas into benign, well differentiated, and moderately or poorly differentiated. The rate of agreement between pathological evaluation and computer classification was approximately 70% in low-grade lesions and 20% in high-grade lesions. These were similar to our results, and the large difference between the results for low-grade and high-grade lesions was also explained by the nuclear variability of breast carcinoma.[8]

The present system was able to divide intraductal lesions into malignant and benign types with a high diagnostic accuracy, especially in well differentiated adenocarcinoma and benign proliferating lesions. Three PU lesions were classified into the NM group, and these lesions might well have been malignant to begin with.

One pathologist called this lesion malignant. We are now planning to compare the neural network judgement with other objective markers, i.e., proliferating cell nuclear antigen and Ki-67 labeling index, and data for oncogene abnormalities.

The rates of false-positive and false-negative lesions were influenced by the classification criteria. Lesions in which more than 60% of the cells were judged to be malignant were classified into the NM group, and those with less than 40% malignant cells into the NB group. However, the number of false-negative lesions was reduced and that of false-positive lesions was increased by using a lower cut-off value (percentage) on the histogram than the criteria used in the present study. Increasing the number of specimens and applying statistical techniques may solve this problem.

Erler et al. constructed neural networks for the diagnosis of hepatocellular carcinoma using morphometric, densitometric, and combined morphometric and densitometric parameters.[12] The morphometric data generated the best results compared with other sets. In our study, the agreement rates in test B using morphometric parameters only and test C using textural parameters only were lower than in test A using combined morphometric and textural parameters. The difference in results was probably derived from the difference in specimen type and parameters employed. Erler et al. used HE-stained slides collected from four different laboratories, which might have shown variations in staining and section thickness, and the parameters they used included a variety of morphometric and a small number of densitometric parameters. On the other hand, we used Feulgen-stained 5-μm-thick sections prepared by a single technician, and employed a number of textural parameters. Textural parameters are probably correlated with chromatin pattern in routine histology. The usefulness of textural analysis for the study of cervical cells[15] and grading of breast carcinoma cells has been reported previously.[16]

Although the results of the present study are encouraging, the degree of accuracy needs to be improved. It became apparent that a computer-assisted system can accurately classify lesions resembling teaching cases. Therefore, selection of an appropriate teaching set and incorporation of additional morphometric and textural parameters may allow construction of an optimal neural network for benign vs. malignant cell classification of intraductal lesions of the breast.

This study concentrated only on nuclear features; addition of structural parameters and analysis of the data as a group, i.e., mean and standard deviation of the collected data, may increase the diagnostic accuracy. Although much further testing is necessary, the present data suggest that application of computer-assisted image analysis and neural networks to histopathological diagnosis and quality control may become possible.

REFERENCES

1) Azzopardi, J. G., Ahmed, A. and Millis, R. R. Epitheliosis and *in situ* carcinoma; papilloma and papillary carcinoma. *In* "Problems in Breast Pathology," ed. J. L. Bennington, pp. 113–166 (1979). Saunders, Philadelphia, PA.

2) Floyd, C. E. Jr., Lo, J. Y., Yun, A. J., Sullivan, D. C. and Kornguth, P. J. Prediction of breast cancer malignancy using an artificial neural network. *Cancer*, **74**, 2944–2948 (1994).

3) Floyd, C. E. Jr. and Tourassi, G. D. An artificial neural network for lesion detection in SPECT images. *Invest. Radiol.*, **27**, 667–672 (1992).

4) Koss, L. G., Lin, E., Schreiber, K., Elgert, P. and Mango, L. Evaluation of the PAPNET™ cytologic screening system for quality control of cervical smears. *Am. J. Clin. Pathol.*, **101**, 220–229 (1994).

5) Wied, G. L., Dytch, H., Bibbo, M., Bartels, P. H. and Thompson, D. Artificial intelligence-guided analysis of cytologic data. *Anal. Quant. Cytol. Histol.*, **12**, 417–428 (1990).

6) Yu, G. H., Sneige, N., Kidd, L. D., Johnston, D. A. and Katz, R. L. Image analysis-derived morphometric differences in fine needle aspirates of ductal and lobular breast carcinoma. *Anal. Quant. Cytol. Histol.*, **17**, 88–92 (1995).

7) Dytch, H. E. and Wied, G. L. Artificial neural networks and their use in quantitative pathology. *Anal. Quant. Cytol. Histol.*, **12**, 379–393 (1990).

8) Dawson, A. E., Austin, R. E., Jr. and Weinberg, D. S. Nuclear grading of breast carcinoma by image analysis: classification by multivariate and neural network analysis. *Am. J. Clin. Pathol.*, **95**, S29–37 (1991).

9) Dybowski, R. and Gant, V. Artificial neural networks in pathology and medical laboratories. *Lancet*, **346**, 1203–1207 (1995).

10) Wolberg, W. H., Street, W. N. and Mangasarian, O. L. Image analysis and machine learning applied to breast cancer diagnosis and prognosis. *Anal. Quant. Cytol. Histol.*, **17**, 77–87 (1995).

11) O'Leary, T. J., Mikel, U. V. and Becker, R. L. Computer-assisted image interpretation: use of a neural network to differentiate tubular carcinoma from sclerosing adenosis. *Mod. Pathol.*, **5**, 402–405 (1992).

12) Erler, B. S., Hsu, L., Truong, H. M., Petrovic, L. M., Kim, S. S., Huh, M. H., Ferrell, L. D., Thung, S. N., Geller, S. A. and Marchevsky, A. M. Image analysis and diagnos-

tic classification of hepatocellular carcinoma using neural networks and multivariate discriminant functions. *Lab. Invest.,* **71,** 446–451 (1994).

13) Dytch, H. E. and Wied, G. L. Neurocomputing: artificial neural nets as AI tools for quantitative pathology. *In* "Manual of Quantitative Pathology in Cancer Diagnosis and Prognosis," ed. J. P. A. Baak, pp. 485–502 (1991). Springer-Verlag, Berlin, Heidelberg.

14) Bacus, J. W. and Bacus, J. V. A method of correcting DNA ploidy measurements in tissue sections. *Mod. Pathol.,* **7,** 652–664 (1994).

15) Pressman, N. J. Markovian analysis of cervical cell images. *J. Histochem. Cytochem.,* **24,** 138–144 (1976).

16) Giménez-Mas, J. A., Sanz-Moncasi, M. P., Remón, L., Gambó, P. and Gallego-Calvo, M. P. Automated textural analysis of nuclear chromatin. A mathematical morphology approach. *Anal. Quant. Cytol. Histol.,* **17,** 39–47 (1995).