Diagnostic Performance of Artificial Neural Network for Detecting Ischemia in Myocardial Perfusion Imaging

Kenichi Nakajima, MD, PhD; Shinro Matsuo, MD, PhD; Hiroshi Wakabayashi, MD, PhD; Kunihiro Yokoyama, MD, PhD; Hisashi Bunko, MD, PhD; Koichi Okuda, PhD; Seigo Kinuya, MD, PhD; Karin Nyström; Lars Edenbrandt, MD, PhD

Background: The purpose of this study was to apply an artificial neural network (ANN) in patients with coronary artery disease (CAD) and to characterize its diagnostic ability compared with conventional visual and quantitative methods in myocardial perfusion imaging (MPI).

Methods and Results: A total of 106 patients with CAD were studied with MPI, including multiple vessel disease (49%), history of myocardial infarction (27%) and coronary intervention (30%). The ANN detected abnormal areas with a probability of stress defect and ischemia. The consensus diagnosis based on expert interpretation and coronary stenosis was used as the gold standard. The left ventricular ANN value was higher in the stress-defect group than in the no-defect group (0.92±0.11 vs. 0.25±0.32, P<0.0001) and higher in the ischemia group than in the no-ischemia group (0.70±0.40 vs. 0.004±0.032, P<0.0001). Receiver-operating characteristics curve analysis showed comparable diagnostic accuracy between ANN and the scoring methods (0.971 vs. 0.980 for stress defect, and 0.882 vs. 0.937 for ischemia, both P=NS). The relationship between the ANN and defect scores was non-linear, with the ANN rapidly increased in ranges of summed stress score of 2–7 and summed defect score of 2–4.

Conclusions: Although the diagnostic ability of ANN was similar to that of conventional scoring methods, the ANN could provide a different viewpoint for judging abnormality, and thus is a promising method for evaluating abnormality in MPI. (Circ J 2015; 79: 1549–1556)

Key Words: Artificial neural network; Coronary artery disease; Myocardial ischemia; Myocardial perfusion imaging; Quantitative analysis

The role of nuclear medicine in cardiology practice has been detection of induced ischemia with respect to diagnosis and optimal treatment strategy. In particular, identification of physiological ischemia with exercise or limited flow reserve by vasodilator stress has been a unique feature of nuclear cardiology using myocardial perfusion imaging (MPI). In the initial stages of nuclear cardiac imaging, the purpose was to detect ischemia in patients with unknown coronary artery disease (CAD), leading to coronary angiography (CAG) and intervention. However, today there are many subjects with multiple sites of coronary stenosis, which may potentially induce ischemia, and modification of ischemia by medical treatment and coronary interventions should also be examined to evaluate hemodynamic changes of the coronary flow reserve. Important indications of MPI included determination of the possibilities of ischemia in the distal site of stenosis, residual ischemia after coronary intervention, and restenosis in the follow-up studies. Therefore, classical diagnostic criteria for detection of abnormality using coronary stenosis as the gold standard (eg, ≥75% or 50%) are apparently insufficient.

Although visual evaluation of the perfusion defect and ischemia is the first step in image interpretation, a semiquantitative approach using a 17-segment model has been most commonly utilized. The most widely used quantification has been the summed stress score (SSS), summed rest score (SRS), and summed differences score (SDS), which reflect the amount of ischemia and infarction. These perfusion abnormalities are still the best predictors of cardiac events even when they are compared with a complete diagnostic work-up and with more recent prognostic indicators. For quantification using scoring, statistical average and deviation of each segment have been...
used, and the Japanese Society of Nuclear Medicine working group databases were created to provide common diagnostic criteria. Recently, artificial intelligence, such as the artificial neural network (ANN), has been used as a new approach in nuclear medicine. Unlike the scoring methods, artificial intelligence mimics and learns interpretation of the experts, and diagnostic suggestion is made by probability of abnormality.

The ANN approach is still in its initial stage, and few studies have used ANN for clinical cardiology practice. The scorers of the defects.

The scoring was based on a threshold method using the average segmental counts and variations. The best-fitted thresholds were decided upon by incorporating optimal maximum-count normalization and subsegmental scores using a Japanese patient group at Kanazawa University. The scoring method in this study was compared with the conventional scoring method using QPS software (Cedars Sinai Medical Center, Los Angeles, CA, USA).

LV functional evaluation was performed, including ejection fraction (EF), and the end-diastolic and end-systolic volumes. The edge-detection algorithm used an active shape model for delineating ventricular edges and a small heart algorithm fitted for Japanese subjects.

ANN for Detecting Ischemia

The first step of the analysis was segmentation of the LV, which was performed using a 3D heart-shaped model and the short-axis slice images. Detection of ischemia was performed in 2 ways. The area of possible perfusion abnormality in the stress images (stress defect) was segmented using a method.

Table 1. Demographics of Patients With CAD Studied With Myocardial Perfusion Imaging

<table>
<thead>
<tr>
<th></th>
<th>Mean±SD (range), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>70±10 (44–87)</td>
</tr>
<tr>
<td>Male sex</td>
<td>65 (61%)</td>
</tr>
<tr>
<td>Height, weight (male)</td>
<td>165±7 cm, 65±11 kg</td>
</tr>
<tr>
<td>Body mass index (male)</td>
<td>24.0±3.3 kg/cm²</td>
</tr>
<tr>
<td>Height, weight (female)</td>
<td>152±7 cm, 53±8 kg</td>
</tr>
<tr>
<td>Body mass index (female)</td>
<td>23.4±3.8 kg/cm²</td>
</tr>
<tr>
<td>Exercise stress</td>
<td>93 (89%)</td>
</tr>
<tr>
<td>No. of vessels (0,1,2,3VD)</td>
<td></td>
</tr>
<tr>
<td>≥75% stenosis</td>
<td>35:35:27:9 (MVD 34%)</td>
</tr>
<tr>
<td>≥50% stenosis</td>
<td>25:29:30:22 (MVD 49%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>61 (58%)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>35 (33%)</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>38 (36%)</td>
</tr>
<tr>
<td>History of MI</td>
<td>29 (27%)</td>
</tr>
<tr>
<td>History of PCI/CABG</td>
<td>30 (28%), 2 (2%)</td>
</tr>
<tr>
<td>ANN value of stress defect</td>
<td>0.65±0.40 (0–1)</td>
</tr>
<tr>
<td>Consensus of stress defect</td>
<td>74 (70%)</td>
</tr>
<tr>
<td>ANN value of ischemia</td>
<td>0.35±0.45 (0–1)</td>
</tr>
<tr>
<td>Consensus of ischemia</td>
<td>39 (37%)</td>
</tr>
<tr>
<td>SSS</td>
<td>8.0±8.7 (0–34)</td>
</tr>
<tr>
<td>SRS</td>
<td>5.5±7.5 (0–35)</td>
</tr>
<tr>
<td>SDS</td>
<td>3.2±3.6 (0–17)</td>
</tr>
<tr>
<td>End-diastolic volume (ml)</td>
<td>104±27 (58–186)</td>
</tr>
<tr>
<td>Ejection fraction (%)</td>
<td>68.8±9.1 (38–86)</td>
</tr>
</tbody>
</table>

ANN, artificial neural network; CABG, coronary artery bypass grafting; MI, myocardial infarction; MVD, multivessel disease; PCI, percutaneous coronary intervention; SDS, summed differences score; SRS, summed rest score; SSS, summed stress score; VD, vessel disease.
mimicking the analysis of physicians who perform visual interpretation of defects. The subtraction image of stress and rest images was created to detect and localize stress-induced ischemia. Features judging the possible perfusion abnormality were analyzed, and included extent, shape, localization, and comparison of count levels with normal stress perfusion databases. The features describing the size and severity of stress defects were calculated in the 3D segmentation of the LV and not in the polar maps in order to avoid the geometrical distortion of these plots. These features were used as input for the ANN system. The training group for SPECT-MPI cases included normal and perfusion defects and was classified by experienced physicians, in which 1,051 patients (498 male and 553 female subjects, 62±10 years, range 29–89 years) were included. In the final output, the ANN value indicated the probability of abnormality; namely, abnormalities of stress defect and ischemia based on stress image and subtraction images, respectively.

Reference of Ischemia and Abnormality in Stress Condition
Final visual judgment of stress abnormality and ischemia was based on the consensus of ≥3 experienced nuclear medicine physicians. In the first evaluation, original short-axis images and polar map were presented with only information of age and sex. Then LV function, including volumes and EF, was added, and subsequently all information of the CAG stenosis, presence of restenosis, and location of stents or bypass grafts was added. The judgment was made in order to classify the information into 4 grades: definitely normal, probably normal, probably abnormal, and definitely abnormal. Existence of stress abnormality and stress-induced ischemia was thus based on an integrated understanding of the coronary stenosis and existence of MI, which was similar to clinical diagnostic procedures for the existence of ischemia.

Apart from the expert visual evaluation, diagnostic accuracy based on coronary stenosis ≥50% as the diagnostic gold standard was also examined in a group of patients without MI and PCI (n=53). In a subgroup of patients with 1-vessel disease who had neither MI nor PCI, detectability of abnormality based on consensus, defect score, and ANN was compared.

Statistical Analysis
All values are expressed as mean±standard deviation (SD). The difference among groups was examined by 1-way analysis of variance with F test and t-test. Nonparametric analysis in multiple comparisons using the Wilcoxon method was also utilized to support the significant differences for each pair. Correlation coefficient and linear regression lines were calculated to compare scores between software programs. Receiver-operating characteristic (ROC) analysis was performed and the area under the curve (AUC) was calculated to evaluate diagnostic ability. A P value <0.05 was considered significant. Statistical analysis was performed using JMP 10.0.2 software (SAS Institute Inc, Cary, NC, USA).

Results
Evaluation of Scores in Comparison With QPS
When the SSS of cardioREPO (y) was compared with that of QPS (x), good correlation was observed between both methods; y=0.099+1.059*x (R²=0.86, P<0.0001). When the SRS and SDS with cardioREPO (y) were compared with those with QPS (x), the correlation of both methods was also good: y=0.465±1.109*x (R²=0.84, P<0.0001) and y=0.458±0.927*x (R²=0.71, P<0.0001), respectively. Classification into 2 groups of SSS ≥8 and SSS <8 resulted in complete agreement in 81 (76%) of 106 patients. Similarly, classification into 2 groups of SDS ≥4 and SDS <4 showed complete agreement in 87 (82%) of the patients.

Consensus for Stress Defect
Regarding the stress defect, groups with definitely normal, probably normal, probably abnormal, and definitely abnormal groups were compared for ANN and SSS values (Table 2A). The ANN values differed significantly among the 4 groups (F ratio 284, P<0.0001), and the SSS also differed significantly (F ratio 59, P<0.0001). Figure 1 shows the results when patients were classified as normal (n=42) or abnormal (n=64). Regarding the ANN value, the normal and abnormal groups showed 0.25±0.32 and 0.92±0.11 (F ratio 237, P<0.0001) and for the SSS value, the 2 groups showed 1.1±1.4 and 12.5±8.6 (F ratio 73, P<0.0001), respectively.

With respect to stress-induced ischemia detected by the subtraction images, groups with definitely normal, probably normal, probably ischemia, and definitely ischemia were compared (Table 2B). Both the ANN value (F ratio 56, P<0.0001) and the SDS (F ratio 34, P<0.0001) differed significantly among the groups. Figure 2 shows results when the patients were classified as normal (n=54) or abnormal (n=52). The normal and abnormal groups showed ANN values of 0.004±0.033 and 0.70±0.40 (F ratio 159, P<0.0001) and for the SDS values, the 2 groups showed 0.80±1.02 and 6.0±3.7 (F ratio 84, P<0.0001), respectively.

The relationship between the ANN and stress defect is shown in Figure 3. The ANN value increased steeply from SSS=2–7 and reached plateau when SSS was ≥8. However, 55 data

<p>| Table 2. ANN Values and Defect Scores in 4 Groups of Expert Visual Evaluation of Ischemia |
|---------------------------------------------|---------|-----------------|-------------------------|</p>
<table>
<thead>
<tr>
<th>Expert interpretation</th>
<th>n</th>
<th>ANN (mean±SD)</th>
<th>Defect score*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Stress defect</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definitely normal</td>
<td>29</td>
<td>0.064±0.11</td>
<td>0.48±0.83</td>
</tr>
<tr>
<td>Probably normal</td>
<td>13</td>
<td>0.65±0.27</td>
<td>2.4±1.4</td>
</tr>
<tr>
<td>Probably abnormal</td>
<td>30</td>
<td>0.86±0.14</td>
<td>7.0±4.5</td>
</tr>
<tr>
<td>Definitely abnormal</td>
<td>40</td>
<td>0.98±0.03</td>
<td>17.4±8.4</td>
</tr>
<tr>
<td><strong>B. Stress-induced ischemia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definitely normal</td>
<td>39</td>
<td>0.0±0.0</td>
<td>0.36±0.58</td>
</tr>
<tr>
<td>Probably normal</td>
<td>15</td>
<td>0.016±0.062</td>
<td>1.9±1.0</td>
</tr>
<tr>
<td>Probably abnormal</td>
<td>21</td>
<td>0.59±0.41</td>
<td>4.4±2.6</td>
</tr>
<tr>
<td>Definitely abnormal</td>
<td>31</td>
<td>0.78±0.39</td>
<td>6.4±4.1</td>
</tr>
</tbody>
</table>

*Defect scores are SSS for stress defect and SDS for stress-induced ischemia. Abbreviations as in Table 1.
of results based on consensus and the ANN or defect scores. In groups of agreement of no stress defect (n=32) and stress defect (n=65), SSS was 0.81±1.40 and 12.4±8.6 (P<0.0001), respectively. However, the discrepant group of positive ANN with negative consensus showed SSS of 1.8±0.8, and the SSS was comparable to that for the group of agreement in the no-ischemia group. In groups of agreement of no-ischemia (n=54) and ischemia (n=39), SDS was 0.80±1.02 and 6.4±3.8, while the discrepant group of the negative ANN with positive consensus showed SDS of 3.1±1.9.

An additional analysis was performed based on conven-

![Figure 1](image1.png)

**Figure 1.** Artificial neural network (ANN) values for stress defect (A) and summed stress score (SSS) (B) for consensus interpretation of perfusion defect under stress conditions. The normal and stress-defect groups differed significantly for ANN (F ratio=237, P<0.0001) and SSS (F ratio=73, P<0.0001). Circles and squares indicate stress defect and no stress defect, respectively. Solid and open marks indicate ischemia and no-ischemia, respectively. Green and blue lines are mean and SD, respectively. Outlier box plot indicates median, 25%, and 75% quartile with whiskers for both ends.

![Figure 2](image2.png)

**Figure 2.** Artificial neural network (ANN) values for ischemia (A) and summed difference score (B) for consensus interpretation of ischemia by subtraction images. The no-ischemia and ischemia groups differed significantly for ANN (F ratio=159, P<0.0001) and SSS (F ratio=84, P<0.0001). Marks are the same as in Figure 1.

points were clustered on SDS=0–3 and ANN=0.0, while high ANN values >0.8 were distributed in a range of SDS ≥4. When QPS and cardioREPO were compared for the 2 groups of SSS ≥8 and <8, agreement of scores between both software programs was obtained in 91 of 106 (86%) patients, including 57 (54%) for SSS <8, and 34 (32%) for SSS ≥8. Similarly, when the 2 groups of SDS ≥4 and <4 were compared, complete agreement of scores was observed in 87 patients (82%), including 61 (58%) for SDS <4 and 26 (25%) for SDS ≥4.

Disagreement between consensus reading and ANN value was examined. Figure 4 shows the relationship among groups
Artificial Neural Network for Myocardial Ischemia

Artificial Neural Network for Myocardial Ischemia (P=NS). Thus, the ANN approach showed comparable diagnostic accuracy to conventional methods using coronary stenosis as the gold standard.

ROC analysis was performed for stress defect and ischemia using the consensus agreement as the gold standard. The ROC AUC for stress defect was 0.980 by SSS (P=NS), whereas it was 0.901 by ANN (P=NS). In addition, when the QPS software was used, ROC AUC was 0.914 by SDS.

Figure 6 shows a patient with 1-vessel disease and how the defect scoring and ANN system interpreted the images for the detected ischemic area.
NAKAJIMA K et al.

Discussion
Artificial intelligence (ie, ANN) was applied to identify myocardial ischemia in this study, and the results were compared with a conventional scoring method using expert reading as the gold standard. The results showed that the ANN provided comparable diagnostic accuracy to the conventional method. The characteristics of judging abnormality by the ANN, however, were quite different from the conventional threshold-based scoring method.

Major difference among visual evaluation, scoring method, and the ANN system should be considered. The scoring method has been based on statistics of mean and deviation in each myocardial segment or pixel-based regional counts. The threshold of abnormality has been empirically determined, because the original scoring method using SSS, SRS, and SDS is based on visual interpretation. In other words, the threshold was adjusted so that similar results were obtained from visual scores. As Garcia et al discuss on Gorry’s reference in their review article, statistical programs cannot “understand” their problem areas, cannot “discuss” their knowledge with the user and have no means of “explaining” it to physicians. The ANN system may potentially offer a natural environment similar to the judgment provided by well-trained physicians, but further progress is required to reach it a level of human expertise.

In the main analysis, the gold standard of the judgment by the ANN was not significant coronary artery stenosis ≥50% or ≥75%, which has been used as the basis for diagnostic ability.

![Figure 5.](image1)

**Figure 5.** Diagnostic accuracy of the expert visual consensus, defect score, and ANN methods when ≥50% coronary stenosis was used as the gold standard. No significant differences in diagnostic accuracy were observed among the 3 diagnostic methods. ANN, artificial neural network.

![Figure 6.](image2)

**Figure 6.** Comparison of abnormality by the defect score and artificial neural network (ANN) methods. Exercise and rest perfusion images are shown for an 87-year-old man with angina pectoris. The locations of stenosis were 99% stenosis of segment 7 and 90% stenosis of segment 9 in the left anterior descending coronary artery. Consensus interpretation was anteroseptal to apical stress-induced ischemia. The defect scores of SSS, SRS, and SDS were 22, 10, and 12, respectively. The ANN system detected abnormality in the anteroseptal to apical regions in the stress (a) (black contour; area 32%, probability 100%) and subtraction (b) (white contour; area 21%, probability 100%) images.
Artificial Neural Network for Myocardial Ischemia

The diagnostic ability might be potentially influenced by the training databases. Johansson et al. presented their initial evaluation of the EXINI heart using the ANN system in comparison with PERFEX (Emory cardiac Toolbox). They reported high diagnostic accuracy with EXINI’s ANN system with 70% sensitivity and 93% specificity to find ischemia, and 84% sensitivity and 84% specificity to find the abnormal study. They included a broad spectrum of diseases, with disease prevalence of 30% according to their clinical routine. On the other hand, our study included patients who underwent CAG showing higher risk for myocardial ischemia because of multivessel diseases. Although detection of ischemia in our study seemed to underestimate the degree of ischemia, the patient selection bias might have influenced the detectability. The training databases are an important factor of the ANN system. Although the ANN system was trained using Swedish databases and based on Swedish expert interpretation of images, comparable diagnostic accuracy was obtained in this Japanese study, probably because the detection of ischemia was based on differences between stress and rest. However, to fit better with Japanese clinical practice and our daily interpretation of the existence of ischemia, retraining of the ANN system might be indicated with a large number of patient groups. Comparable retraining of the neural network system using Japanese databases successfully enhanced diagnostic ability in bone scan interpretation using the ANN.

Study Limitations
This study included complicated patients, including multivessel disease, history of MI, and coronary revascularization. However, the diagnostic ability regarding significant stenosis (eg., ≥50% stenosis) as the gold standard, also showed comparable diagnostic ability between the experts and the ANN. Approximately 30% of the patients already had MI, and the mean SDS was 3.2±3.6, the score corresponding to pure 10% ischemia could not be defined. As discussed, retraining of the ANN system with Japanese databases may enhance the diagnostic ability if a new project including a larger number of patients, in the order of >1,000 patients, is designed. Location of the coronary artery territory was not included for the training of the ANN system at present. Although it could be included in the training process, strict correspondence should be performed with coronary computed tomography and MPI fusion imaging. Finally, clinical information and LV function could be included as input features, which is a comparable situation to clinical decision-making.

Conclusions
The neural network system was applied to the diagnosis of stress defect and ischemia in MPI. The diagnostic ability was comparable to expert visual interpretation and conventional scoring methods. Even when significant coronary stenosis was used as the gold standard in patients without MI and revascularization, the diagnostic ability of the ANN was comparable to that of the conventional methods. The ANN value provided probability of abnormality, mimicking the process of human interpretation, and the ANN system could provide diagnostic suggestions from a different viewpoint than the statistical scoring method. We conclude that the ANN system could be a promising new adjunctive method for the detection of myocardial ischemia.
Acknowledgments

The authors thank the technologists of Kanazawa University Hospital, Matto Ishikawa Central Hospital, and Kanazawa Cardiovascular Hospital for their cooperation in accumulating the DICOM data. We also thank Mr Ronald Belisle for his editorial assistance.

Disclosures

Conflict of interest: K. Nakajima has a collaborative research work with FIJIFILM RI Pharma Co Ltd (Tokyo, Japan) for the development of software programs. K. Nyström and L.E. are employed by and are shareholders of EXINI Diagnostics AB (Lund, Sweden).

References


