

Usefulness of an artificial neural network for a beginner to achieve similar interpretations to an expert when examining myocardial perfusion images

Akiyo Chiba (✉ akiyopoo123@icloud.com)

Nagasaki University Graduate School of Biomedical Sciences <https://orcid.org/0000-0002-6887-413X>

Takashi Kudo

Department of Radioisotope Medicine, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki

Reiko Ideguchi

Department of Radioisotope Medicine, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki

Altay Myssaev

Department of Radioisotope Medicine, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki

Seiji Koga

Department of Cardiovascular Medicine, Nagasaki University Hospital, Nagasaki

Tsuyoshi Yonekura

Department of Cardiovascular Medicine, Nagasaki University Hospital, Nagasaki

Akira Tsuneto

Department of Cardiovascular Medicine, Nagasaki University Hospital, Nagasaki

Minoru Morikawa

Department of Radiology, Nagasaki University Hospital, Nagasaki

Satoshi Ikeda

Department of Cardiovascular Medicine, Nagasaki University Hospital, Nagasaki

Hiroaki Kawano

Department of Cardiovascular Medicine, Nagasaki University Hospital, Nagasaki

Yuji Koide

Nagasaki Medical Education Center, Nagasaki University Hospital, Nagasaki

Masataka Uetani

Department of Radiological Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki

Koji Maemura

Department of Cardiovascular Medicine, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki

Original research

Keywords: artificial neural network, myocardial perfusion images, artificial intelligence, nuclear medicine, cardiology imaging

Posted Date: September 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-78078/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at The International Journal of Cardiovascular Imaging on March 11th, 2021. See the published version at <https://doi.org/10.1007/s10554-021-02209-z>.

Abstract

Background: This study examined whether using an artificial neural network (ANN) helps beginners in diagnostic cardiac imaging to achieve similar results to experts when interpreting stress myocardial perfusion imaging (MPI).

Methods: 138 patients underwent stress MPI with Tc-labeled agents. An expert and a beginner interpreted stress/rest MPI with or without the ANN and the results were compared. The myocardium was divided into 5 regions (the apex; septum; anterior; lateral, and inferior regions), and the degree of myocardial blood flow was evaluated from 0 to 4, and SSS, SRS, and SDS were calculated. The ANN effect, defined as the difference in each of these scores between with and without the ANN, was calculated to investigate the influence of ANN on the interpreters' performance.

Results: The ANN effect was smaller for the expert than for the beginner.

Conclusions: When interpreting MPI, beginners may achieve similar results to experts by using an ANN. Thus, interpreting MPI with ANN may be useful for beginners.

Introduction

During the last 20 years, artificial intelligence (AI) and machine learning (ML), including artificial neural networks (ANN), have markedly developed. AI and ML may aid medical imaging-based diagnosis, not only in terms of the detection of disease, but also in management, reporting, and prognostication. [1] ANN is a computational model of ML based on the human brain. It has been found that ANN are powerful tools for pattern recognition, signal processing, image or speech data compression, and learning expert systems. [2] There are 6675 radiologists with specialty licenses in Japan, according to the Japan Radiological Society, but only 1317 radiologists have specialty licenses in nuclear medicine. In particular, the number of experts in cardiac nuclear medicine is very small. The interpretation of myocardial perfusion images (MPI) requires skilled expert reading, but there are not as many skilled experts in the interpretation of MPI as there are in the interpretation of computed tomography scans in Japan.

Recently, many studies have shown that AI and ML significantly improve the diagnostic accuracy not only in cardiology, but also in many other medical fields. [3, 4] However, there has not been sufficient research on the use of AI in cardiac nuclear medicine. Nuclear medicine imaging techniques, such as MPI, do not provide as much data as other imaging techniques; therefore, AI may be useful for inexperienced physicians working in radiology or cardiology when they interpret MPI. If the use of AI helps beginners to achieve similar interpretations to experts when examining MPI, it would help to compensate for a lack of human resources.

The purpose of this study was to examine whether a beginner in diagnostic cardiac imaging can achieve similar results to an expert when interpreting stress MPI by using an ANN.

Methods

Subjects

The subjects were 138 consecutive patients who underwent stress MPI (with Tc-labeled agents) at Nagasaki University Hospital between May 2014 and June 2015, including 52 patients (38%) with multivessel disease, 35 patients (25%) with a history of myocardial infarction, and 21 patients (15%) with both multivessel disease and a history of myocardial infarction. Of these 138 patients, 8 patients (6%) underwent coronary CT scan, 61 patients (44%) underwent coronary angiography, and 26 patients (19%) underwent both coronary CT scan and coronary angiography. Of 52 patients with multi vessel disease, 1 patient (2%) underwent coronary CT scan, 32 patients (62%) underwent coronary angiography, and 19 patients (37%) underwent both of coronary CT scan and coronary angiography. An expert with over 20 years' interpretation experience and a beginner with a few years' interpretation experience interpreted stress MPI with/without software, which implemented ANN, and the results were compared. They interpreted randomly these MPI without any information except patients' age and sex. The patients' characteristics are shown in Table 1.

Table 1
Demographics of patients that underwent stress myocardial perfusion imaging

n = 138	Mean ± SD (range), n (%)
Age (years)	70.6 ± 0.8 (36–87)
Sex (male)	94 (68.1%)
Height, weight (male)	164 ± 0.7 cm, 61.3 ± 1.2 kg
Body mass index (male)	22.7 ± 0.4 kg/cm ²
Height, weight (female)	149.7 ± 1.0 cm, 48.8 ± 1.5 kg
Body mass index (female)	21.8 ± 0.6 kg/cm ²
No. of vessels displaying	
≥75% stenosis (0, 1, 2, 3)	63:23:17:35 (MVD: 38%)
Hypertension	94 (68.1%)
Diabetes mellitus	53 (38.4%)
Dyslipidemia	93 (67.4%)
History of MI	35 (25.3%)
History of PCI/CABG	53 (38.4%), 18 (13%)
LVEF (%) (QGS, stress)	67.4 ± 1.4
LVEF (%) (QGS, rest)	68.7 ± 1.0
LVEDV (ml) (QGS, stress)	31.8 ± 3.3
LVEDV (ml) (QGS, rest)	32.5 ± 2.2
CABG, coronary artery bypass grafting; MI, myocardial infarction;	
MVD, multivessel disease; PCI, percutaneous coronary intervention; LVEF, left ventricle ejection fraction; LVEDV, left ventricle end-diastolic volume; QGS, quantitative gated SPECT	

Imaging

The stress and rest MPI studies were performed using a 1-day stress-first protocol and about 1200 MBq of a ^{99m}Tc-labeled myocardial perfusion agent (tetrofosmin or MIBI; divided into 300 MBq for the stress imaging and 900 MBq for the rest imaging). The indications and stress protocols followed the guidelines of the Japanese Circulation Society [5]. An adenosine stress test was performed with a standard continuous injection protocol, involving an injection rate of 0.12 mg•kg⁻¹•min⁻¹, in 136 patients (98.6%), and an exercise stress test, involving symptom-limited ergometer exercise, was conducted in 2 patients

(1.4%). The end-points of the exercise stress test included significant symptoms (such as chest pain, dyspnea, or leg fatigue), the achievement of the target heart rate, electrocardiographic changes (ST depression, ST elevation, fatal arrhythmia, or blood pressure problems [very high pressure of > 250 mmHg or hypotension]). One hour after the injection of the tracer, the patients were imaged using a dual-headed single-photon emission computed tomography (SPECT) system, equipped with low-energy high-resolution collimators (e.cam Signature; Siemens Healthcare GmbH, Germany); a 180° arc; and a 16 frames/beat acquisition protocol. The acquisition energy level was set at 140 keV with a 20% window fitted for ^{99m}Tc. All of the patients were instructed to refrain from eating food (breakfast) before the scans.

Image interpretation and scoring

The images were interpreted and scored in a medical image viewer, using the hospital's Picture Archiving And Communication System (PACS) (Synapse; Fujifilm, Tokyo, Japan) with or without the diagnostic ANN software, (cardioREPO, version 1.1; Fujifilm Toyama Chemical, Tokyo, Japan). This software analyzes MPI via a ML system, which was trained using about 1000 patients' images, and displays areas of abnormal stress perfusion and ischemic areas on a polarmap. The details of the method are described in Fig. 1. [6, 7]

Usually, a 17-segment model is used for myocardial perfusion scoring. [8] However, our study aimed to examine the effects of AI on the performance of beginners during image interpretation. To simplify the scoring for the beginner, the myocardium was divided into 5 regions; i.e., the apex; septum; and the anterior, lateral, and inferior regions. The degree of myocardial blood flow was classified from 0 to 4, and the summed stress score (SSS), summed rest score (SRS), and summed difference score (SDS) were calculated. These parameters were obtained with and without the ANN for both interpreters. In order to investigate the influence of the ANN on the interpreters' performance, the ANN effect was calculated as the difference in each score (SSS, SRS, or SDS) between with and without the ANN for both the beginner and expert. A larger ANN effect indicates that the findings were interpreted more accurately with than without the ANN.

Statistics

All data are expressed as mean ± standard deviation (SD) values. The significance of differences was examined by one-way analysis of variance with the F test or paired t-test. P-values of < 0.05 were considered to be significant. All statistical analyses were performed using the JMP 10.0.2 software.

All procedures performed were in accordance with the ethical standards of the institutional research committee and with the principles of the 1964 Declaration of Helsinki and its later amendments. All clinical data were completely anonymized and processed at Nagasaki University Hospital. This study was approved by Nagasaki University Hospital Clinical Research ethics committee (approval No. 15072762).

Results

Interpretation of images

The results are shown in Table 2. The ANN effect was smaller for the expert than for the beginner (SSS: -0.49 vs. -1.23, $p < 0.0001$; SRS: -0.34 vs. -0.88, $p = 0.0003$; SDS: -0.15 vs. -0.36, $p = 0.0128$, respectively). The ANN effects on SSS, SRS, and SDS were negative for both interpreters, which indicates that they had lower scores with than without the ANN; i.e., their diagnostic approach became more conservative when they were using the ANN. The absolute ANN effect value was lower in the expert than in the beginner, which means that the abovementioned effect was larger in the beginner. When the scores obtained by the expert with the ANN were defined as standard scores, the SSS and SRS scores of the beginner were closer to the standard scores with than without the ANN.

Table 2
Results

ANN effect	Expert	Beginner	p-value
SSS	-0.49 ± 0.08	-1.23 ± 0.15	< 0.0001
SRS	-0.34 ± 0.07	-0.88 ± 0.13	0.0003
SDS	-0.15 ± 0.06	-0.36 ± 0.08	0.0128
ANN, artificial neural network; SSS, summed stress score; SRS, summed rest score; SDS-summed difference score			

Discussion

Our results indicate that an ANN reduced the image interpretation gap between an expert and beginner. In a previous study, neural network software showed sensitivity of 90% and specificity of 85% for detecting myocardial ischemia, which were superior to the sensitivity and specificity of a computer-assisted diagnostic system (the Emory Cardiac Toolbox). In addition, a decision support system based on neural networks achieved interpretations that were more similar to those of experienced clinicians than those achieved by a conventional automated quantification software package. [9] Our study did not compare multiple methods, but rather evaluated the additional effect of using AI to aid image interpretation. However, both studies showed that AI improves the interpretation of MPI.

Our results regarding the skill effect showed that the SDS scores of the beginner were lower than those of the expert when the ANN was used. If the expert's scores are taken as standard scores, this finding indicates that the beginner made false-negative interpretations of ischemia when using the ANN. In a study by Nakajima, [10] when the ANN threshold was set at 0.5, the ANN interpreted regions without ischemia perfectly. However, it missed some cases that were interpreted as ischemia by the consensus reading. This was especially common in cases of mild ischemia or mild infarctions. These findings indicate that compared with the expert consensus the ANN produced false-negatives for ischemia, which may explain our results. If a beginner interpreted images as suggested by such an ANN, but an expert was

able to interpret them correctly, the beginner would produce false-negatives for ischemia when using the ANN. However, the interpretations of a single expert will not always be correct. Our results showed that even experts can change their interpretations of images when using ANN although the magnitude of the changes made by the expert was smaller than that of the changes made by the beginner. In a previous study, images were interpreted using an ANN and re-evaluated by 3 experts, and differences were found in 53 of 200 cases. This disagreement was related to small or mild perfusion defects, which indicates that the interpretations of the experts may have fluctuated or even been wrong.^[11] An ANN may be useful for standardizing image interpretation, even for experts, which may explain why the scoring by the expert was still affected by the use of the ANN.

We used an ANN to aid image analysis/interpretation. Arsanjani et al. [12] showed that the LogitBoost method, which is another type of ML, exhibited almost the same accuracy when interpreting MPI as expert readers. Our findings are similar.

Limitations

This study involved an expert and a beginner in cardiac nuclear medicine and was performed at Nagasaki University. Although, we should have included a few experts and a few beginners to obtain sufficient data, there was only one expert in cardiac nuclear medicine in Nagasaki Prefecture.

In addition, we used a simplified 5-segment model instead of the conventional 17-segment model for scoring. This narrowed the range of scores (for 5 segments, the maximum value is 20, whereas for 17 segments the maximum value is 68), which may have reduced the statistical power of the study. However, a previous study, which analyzed a neural network support system, also used a 5-segment model. Thus, we consider that this 5-segment model did not markedly affect our results [9].

Strictly speaking, we should compare our study results with gold standard, but our study target is to achieve going beginners' interpretation with AI up to human experts' that. So we didn't do it in this study. [7]

Conclusions

When using an ANN, beginners may achieve similar imaging interpretations to experts in cardiac nuclear medicine. Furthermore, ANN systems may be useful for obtaining a second opinion, particularly when physicians are inexperienced at interpreting nuclear cardiology imaging.

Abbreviations

MPI: myocardial perfusion imaging; AI: artificial intelligence; ANN: artificial neural network; SSS: summed stress score; SRS: summed rest score; SDS: summed difference score; ML: machine learning

Declarations

Acknowledgments

No applicable

Ethics approval and consent to participate

All procedures performed were in accordance with the ethical standards of the institutional research committee and with the principles of the 1964 Declaration of Helsinki and its later amendments. All clinical data were completely anonymized and processed at Nagasaki University Hospital. This study was approved by Nagasaki University Hospital Clinical Research ethics committee (approval No. 15072762).

Consent for publication

Not applicable

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

All authors declare that they have no conflict of interest.

Funding

No financial support was used for this study.

Author details

¹ Department of Cardiovascular Medicine, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan.

² Department of Radioisotope Medicine, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki, Japan.

³ Department of Radiology, Nagasaki University Hospital, Nagasaki, Japan.

⁴ Department of Cardiovascular Medicine, Nagasaki University Hospital, Nagasaki, Japan.

⁵ Department of Radiological Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan.

⁶ Nagasaki Medical Education Center, Nagasaki University Hospital, Nagasaki, Japan.

⁷ Department of Cardiovascular Medicine, Nagasaki Memorial Hospital, Nagasaki, Japan

References

1. Nensa F, Demircioglu A, Rischpler C. Artificial Intelligence in Nuclear Medicine. *J Nucl Med.* 2019;60(Suppl 2):29s-37s.
2. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. *J Nucl Med.* 1992;33(2):272-6.
3. Hu LH, Betancur J, Sharir T, Einstein AJ, Bokhari S, Fish MB, et al. Machine learning predicts per-vessel early coronary revascularization after fast myocardial perfusion SPECT: results from multicentre REFINE SPECT registry. *Eur Heart J Cardiovasc Imaging.* 2019.
4. Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, et al. A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using (18)F-FDG PET of the Brain. *Radiology.* 2019;290(2):456-64.
5. Guidelines for clinical use of cardiac nuclear medicine (JCS 2010) - digest version. *Circ J.* 2012;76(3):761-7.
6. Nakajima K, Kudo T, Nakata T, Kiso K, Kasai T, Taniguchi Y, et al. Diagnostic accuracy of an artificial neural network compared with statistical quantitation of myocardial perfusion images: a Japanese multicenter study. *Eur J Nucl Med Mol Imaging.* 2017;44(13):2280-9.
7. Nakajima K, Okuda K, Watanabe S, Matsuo S, Kinuya S, Toth K, et al. Artificial neural network retrained to detect myocardial ischemia using a Japanese multicenter database. *Ann Nucl Med.* 2018;32(5):303-10.
8. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *Int J Cardiovasc Imaging.* 2002;18(1):539-42.
9. Lomsky M, Gjertsson P, Johansson L, Richter J, Ohlsson M, Tout D, et al. Evaluation of a decision support system for interpretation of myocardial perfusion gated SPECT. *Eur J Nucl Med Mol Imaging.* 2008;35(8):1523-9.
10. Nakajima K, Matsuo S, Wakabayashi H, Yokoyama K, Bunko H, Okuda K, et al. Diagnostic Performance of Artificial Neural Network for Detecting Ischemia in Myocardial Perfusion Imaging. *Circ J.* 2015;79(7):1549-56.
11. Tagil K, Marving J, Lomsky M, Hesse B, Edenbrandt L. Use of neural networks to improve quality control of interpretations in myocardial perfusion imaging. *Int J Cardiovasc Imaging.* 2008;24(8):841-8.
12. Arsanjani R, Xu Y, Dey D, Vahistha V, Shalev A, Nakanishi R, et al. Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. *J Nucl Cardiol.* 2013;20(4):553-62.

Figures

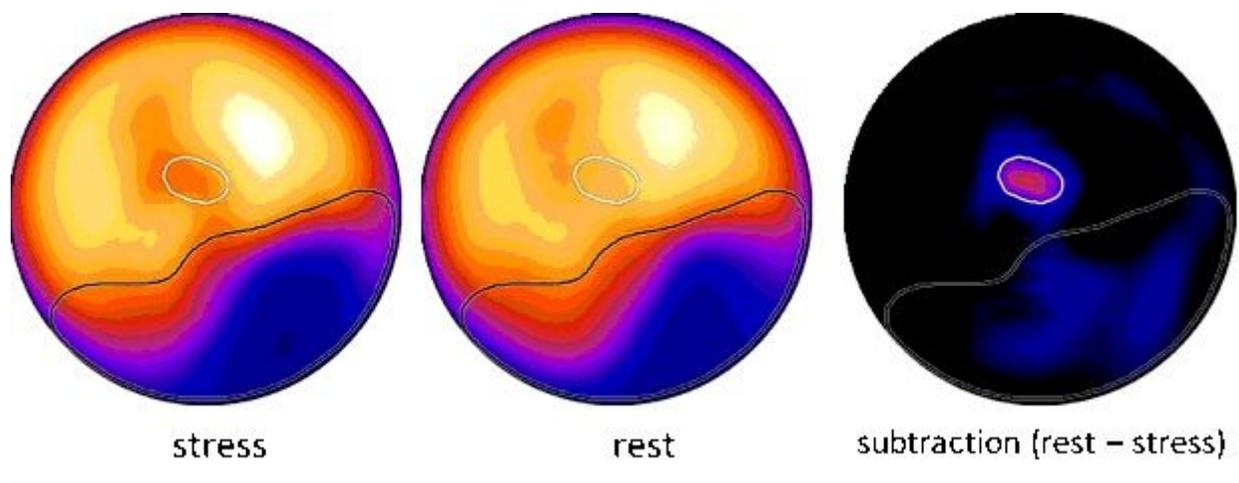


Figure 1

ANN analysis of stress MPI by cardioREPO • Nuclear cardiology experts judged all of the regions on MPI as abnormal (true) or normal (false). • Data from 1001 patients who underwent stress/rest MPI were collected from 12 institutions in Japan. • The final judgments regarding the candidate regions made at the core center served as the gold standard and were used to train the ANN system. • Based on the initial ANN training, differences between visual and ANN findings were selected and used to appropriately adjust the diagnostic thresholds. • ANN value: probability that the lesion is abnormal (values of ≥ 0.5 are significant) • The region within the black line: the region exhibiting abnormal perfusion • The region within the white line: the ischemic region • The region within the black line, but not within the white line: a myocardial infarction Kenichi Nakajima, et al. Eur J Nucl Mol Imaging (2017) 44: 2280–2289