Original Article

Evaluation of myocardial perfusion imaging techniques and artificial intelligence (AI) tools in coronary artery disease (CAD) diagnosis through multi-criteria decision-making method

Hasan Erdagli[1], Dilber Uzun Ozsahin[2,3,4], Berna Uzun[4,5]

[1]Department of Biomedical Engineering, Near East University, Nicosia, Turkey; [2]Department of Medical Diagnostic Imaging, College of Health Science, University of Sharjah, Sharjah, United Arab Emirates; [3]Research Institute for Medical and Health Sciences, University of Sharjah, Sharjah, United Arab Emirates; [4]Operational Research Center in Healthcare, Near East University, Nicosia, Turkey; [5]Department of Mathematics, Near East University, Nicosia, Turkey

Contributions: (I) Conception and design: All authors; (II) Administrative support: D Uzun Ozsahin, B Uzun; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: H Erdagli; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Hasan Erdagli, MSc. Department of Biomedical Engineering, Near East University, Near East Boulevard, TRNC Mersin 10, Nicosia 99138, Turkey. Email: hasan.erdagli@neu.edu.tr.

Background: Cardiovascular diseases (CVDs) continue to be the world's greatest cause of death. To evaluate heart function and diagnose coronary artery disease (CAD), myocardial perfusion imaging (MPI) has become essential. Artificial intelligence (AI) methods have been incorporated into diagnostic methods such as MPI to improve patient outcomes in recent years. This study aims to employ a novel approach to examine how parameters/criteria and performance metrics affect the prioritization of selected MPI techniques and AI tools in CAD diagnosis. Identifying the most effective method in these two interconnected areas will increase the CAD diagnosis rate.

Methods: The study includes an in-depth investigation of popular convolutional neural network (CNN) models, including InceptionV3, VGG16, ResNet50, and DenseNet121, in addition to widely used machine learning (ML) models, including random forests (RF), K-nearest neighbor (KNN), support vector machine (SVM), and Naïve Bayes (NB). In addition, it includes the evaluation of nuclear MPI techniques, including positron emission tomography (PET) and single photon emission computed tomography (SPECT), with the non-nuclear MPI technique of cardiovascular magnetic resonance imaging (CMR). Various performance metrics were used to evaluate AI tools. They are F1-score, recall, specificity, precision, accuracy, and area under the receiver operating characteristic curve (AUC-ROC). For MPI techniques, the evaluation criteria include specificity, sensitivity, radiation dose, cost of scan, and study duration. The analysis was evaluated and compared using the fuzzy-based preference ranking organization method for enrichment evaluation (PROMETHEE), the multi-criteria decision-making method (MDCM).

Results: According to the study's findings, considering selected performance metrics or criteria, RF is the most efficient AI tool for SPECT MPI in the diagnosis of CAD with a net flow (Φ^{net}) of 0.3778, and it's revealed that CMR is the most efficient MPI technique for CAD diagnosis with a net flow of 0.3666. By expanding this study, more comprehensive evaluations can be made in the diagnosis of CAD.

Conclusions: It was concluded that CMR outperformed the nuclear MPI techniques. SPECT, as the least advantageous technique, remained below average on other criteria except for the cost of the scan. Integrating the RF algorithm, which stands out as the most effective AI tool in diagnosing CAD, with SPECT MPI may contribute to SPECT becoming a superior alternative.

Keywords: Coronary artery disease (CAD); diagnostic methods; artificial intelligence tools (AI tools); multi-criteria decision-making; fuzzy preference ranking organization method for enrichment evaluation (fuzzy PROMETHEE)

Submitted May 27, 2024. Accepted for publication Sep 29, 2024. Published online Dec 09, 2024.
doi: 10.21037/cdt-24-237
View this article at: https://dx.doi.org/10.21037/cdt-24-237

Introduction

Background

Cardiovascular disease (CVD) refers to a wide range of conditions impacting the heart and blood vessels. It encompasses a variety of disorders, such as peripheral vascular diseases, valve diseases, heart failure, and coronary artery disease (CAD) (1). Genetic predisposition and modifiable risk factors, such as high blood pressure, increased cholesterol, smoking, obesity, diabetes, and a sedentary lifestyle, frequently combine to cause these illnesses (2). CVDs are the major global source of morbidity and mortality, which has a significant influence on public health (3). The consequences of CVD can be severe, leading to complications such as strokes, heart attacks, and other potentially fatal events. Therefore, knowledge of the mechanisms and risk factors involved is essential for efficient prevention, diagnosis, and treatment of CVD.

A crucial component of cardiac physiology, myocardial perfusion describes how the heart muscle, or myocardium, gets blood supply (4). This blood supply is vital for maintaining the myocardial tissue's metabolic needs by supplying oxygen and other necessary nutrients. Optimal myocardial perfusion is essential for maintaining cardiac function. In cases where blood flow to the myocardium is compromised, a condition known as ischemia may occur, potentially leading to damage to the heart muscle. Coronary artery function, blood pressure, and general cardiovascular health are some variables that affect the dynamic process of myocardial perfusion. Cardiovascular diagnostics include monitoring and assessment of myocardial perfusion, which aid in the identification of abnormalities and the development of suitable treatment plans by medical professionals (4). These diagnostic approaches include techniques such as echocardiography, cardiac computed tomography angiography (CCTA), cardiovascular magnetic resonance imaging (CMR), single photon emission computed tomography (SPECT), and positron emission tomography (PET). Complex medical techniques known as nuclear imaging technology are used to visualize inside body structures and processes within the body (5). These techniques are essential for determining myocardial perfusion in the field of cardiovascular medicine. Specialized imaging techniques, such as PET or SPECT, obtain images of myocardial perfusion after the administration of a radiotracer, which is selectively taken up by the myocardium in proportion to blood flow (6). When stress and rest images are compared, regions with decreased circulation under stress can be identified, which may indicate ischemia. The information, which is associated with the patient's clinical symptoms and history, serves to evaluate the severity, localization, and degree of CAD. In addition, the CMR technique used in CAD is an important tool for myocardial perfusion assessment. CMR provides high-resolution anatomical and functional information. This technique involves the use of contrast material to evaluate myocardial blood flow. After contrast material is injected into the blood vessels, it is used to visualize the blood supply pattern of the myocardium in images taken with a CMR device (7). Compared to nuclear imaging,

Highlight box

Key findings
- Random forests (RF) stands out as the most effective option as an artificial intelligence (AI) tool for myocardial perfusion imaging (MPI) in the diagnosis of coronary artery disease (CAD).
- Cardiovascular magnetic resonance imaging (CMR) stands out as the most effective technique for MPI in the diagnosis of CAD.

What is known and what is new?
- While previous research has explored AI tools and MPI techniques, few alternatives and criteria have been employed, and the importance weights of the criteria were not taken into evaluation.
- Although single photon emission computed tomography (SPECT) is widely used in the diagnosis of CAD among nuclear MPI techniques, non-nuclear CMR has been found to have superior performance. A multi-criteria decision-making method (MCDM) can evaluate the MPI devices' weaknesses in CAD diagnosis. Studies can be conducted to improve the identified deficiencies.

What is the implication, and what should change now?
- The inclusion of the RF algorithm, which has been revealed as the most effective AI tool for the diagnosis of CAD, in SPECT MPI may enable SPECT, the most widely preferred nuclear imaging technique, to rise to first place in the preference rankings. By writing an MCDM application, a more comprehensive decision matrix can be created in CAD diagnosis, and a special preference ranking evaluation can be made for patients.

CMR shows some differences. First, CMR does not involve radiation, making it safe when repeated imaging is required (8). However, in some cases, the application of CMR may be subject to limitations. For example, in patients who are allergic to contrast agents, alternative imaging methods may be preferred. Similarly, the risk of contrast media affecting the kidneys should be considered in individuals with serious kidney problems. In the presence of metal implants, potential risks such as heating or movement due to magnetic fields should be taken into account (8). Therefore, nuclear imaging techniques may provide access to a wider range of patients in some cases. The clinical situation and requirements must be taken into account to determine the most appropriate imaging modality for each patient. In the field of cardiology, nuclear imaging techniques are particularly beneficial due to their non-invasive nature, which has greatly advanced patient care and diagnostics related to CVD (9,10). Even though myocardial perfusion imaging (MPI) is non-invasive, it has limitations, such as the possibility of false-negative and false-positive results (11). However, its value lies in its capacity to offer a thorough evaluation of myocardial perfusion dynamics, directing treatment and diagnostic approaches for CAD patients. Through the application of artificial intelligence (AI), the limitations related to the diagnosis of results derived from MPI can be surmounted (12). During this process, performance metrics such as specificity, accuracy, precision, recall, area under the receiver operating characteristic curve (AUC-ROC), and F1-score, which are of high significance in AI tools, should be considered. The incorporation of AI into MPI is a novel approach in the diagnostic field of CAD. AI systems have proven to be unique in their ability to interpret complex patterns from MPI data, allowing for the more precise and nuanced identification of perfusion anomalies and myocardial ischemia associated with CAD (13). The ability of these AI systems to quickly analyze large datasets and identify tiny subtleties that are invisible to the human eye can improve diagnostic precision. AI and MPI together have a lot of potential to improve diagnostic efficacy and refine limitations (14).

Several AI models, including traditional ML and CNN systems, are employed for the diagnosis of nuclear MPI (15). CNNs are neural networks specifically designed for image processing, and they are exceptionally adept at automatically deriving complex patterns and characteristics from image input. They are highly skilled at capturing complex relationships, which helps them achieve high accuracy when interpreting images. However, CNNs might require a significant amount of processing resources to train (16). On the other hand, while standard ML models serve efficiency and interpretability, they could have trouble identifying subtle patterns found in complicated medical images. The size and kind of the data type, the necessity for interpretability, processing resources, and other considerations all influence the choice between CNNs and ML models (17).

Rationale and knowledge

Various AI approaches for detecting CAD have been compared in previous studies. Papandrianos *et al.* (15) employed and constructed a CNN model for the diagnosis of ischemia or infarction based on SPECT-MPI scans. Furthermore, they conducted a comparative analysis with other CNN models. The outcomes of their study demonstrated that the utilized methods exhibit considerable accuracy and capability in distinguishing between infarction or ischemia and healthy patients. Furthermore, Cantoni *et al.* (18) compared the prognostic values of stress MPI performed with conventional SPECT (C-SPECT) and cadmium-zinc-telluride (CZT) SPECT and revealed the performance differences between these two technologies using machine learning (ML) approaches. The study evaluated the stress MPI results performed using both C-SPECT and CZT-SPECT on 453 patients. The ML tools used in the study include random forests (RF), K-nearest neighbor (KNN), support vector machine (SVM), Naïve Bayes (NB) and decision tree. Accuracy, recall, specificity, and AUC-ROC were calculated as performance metrics. The results show that CZT-SPECT provides higher accuracy and sensitivity than C-SPECT in terms of CAD detection. In particular, it was determined that CZT-SPECT provided significant superiority in terms of accuracy and sensitivity in the analyses performed with KNN and SVM algorithms (P values for accuracy were 0.021 and 0.016, respectively; P values for sensitivity were 0.001 and 0.028, respectively). In addition, a significant improvement was observed in the overall performance of the models after the synthetic minority oversampling technique (SMOTE) was applied to eliminate data imbalance. In particular, it was observed that CZT-SPECT performed better in terms of accuracy, sensitivity, and specificity in the analyses performed with RF, SVM, and KNN algorithms. These findings reveal that CZT-SPECT offers higher performance and better prognostic value. Xu *et al.* (19) compared the efficacy of CMR, SPECT, and PET in the diagnosis of CAD through meta-analysis. Analysis of 203 studies obtained from PubMed, Web of Science, EMBASE, and Cochrane Library databases revealed that the sensitivity of CMR was 0.86, SPECT was 0.83, and PET was 0.85. The specificity of CMR was 0.83, SPECT was 0.77, and PET was 0.86. In particular, CMR and PET were found to have higher diagnostic performance. The study highlights the advantages of CMR and PET in the non-invasive diagnosis of CAD and provides important data for clinical decision support.

Performance evaluation metrics and criteria have been the consistent basis for model evaluation in all of the research papers listed above and many others. However, these studies have not proposed additional critical criteria or classified the importance of the existing criteria to ensure flexible, robust, and comprehensive models. This raises issues that may arise when decision-makers need additional information that is not included in the performance evaluation parameters. What happens if the decision-makers have concerns about how usable the accurate model is? What kind of consequences do changes in the importance of performance metrics lead to? Or how effective the widely used nuclear MPI techniques are in terms of CAD diagnostic success and patient benefits? There are still no responses to any of these questions. As a result, there has become a research deficiency in CAD diagnosis.

Analyzing SPECT-MPI results according to performance metrics using CNN and ML models plays an important role in the process of developing Food and Drug Administration (FDA)-approved software. This analysis helps determine the model with the highest accuracy and lowest error rate, which increases the clinical reliability of the software. The generalizability of the best-performing model, producing consistent results in different patient groups, and being able to adapt to data diversity, increases the wide applicability of the software in the clinical environment. These features ensure that the software complies with legal regulations and can be used safely in clinical environments. Ultimately, this analysis may provide a starting point and reference to develop FDA-approved software by comparing AI tools.

Objective

This study provides an evaluation of AI tools used for MPI to diagnose CAD using a Multi-Criteria Decision Making (MCDM) approach. Frequently used models for CNN, such as VGG16, ResNet50, InceptionV3, and DenseNet121, and commonly applied ML models, including KNN, SVM, NB, and RF, were analyzed. To analyze and compare the AI tools, specificity, accuracy, precision, recall, AUC-ROC, and F1-score values were examined as shown in *Table 1*. SPECT-MPI was referenced when obtaining the data for the comparison and analysis of AI tools. In addition, for MPI, both nuclear imaging techniques (SPECT and PET) and an alternative method, CMR, were compared and analyzed. Although there are various MPI techniques, these three alternatives were preferred for this study due to their wide application areas, high accessibility, and literature gaps. Such a comprehensive examination of nuclear and CMR techniques facilitated by the MCDM process will reveal the strengths and weaknesses of each method and allow for more informed choices in clinical decision-making processes. To analyze and compare the MPI techniques, specificity, sensitivity, radiation dose, cost of scan, and study duration values were examined, as shown in *Table 2*. MCDM, specifically the fuzzy-based preference ranking organization method for enrichment evaluation (PROMETHEE), enables the evaluation of AI tools and MPI techniques using various criteria and performance metrics, in addition to the typically used parameters for diagnosing CAD. This will ensure that decision-makers have access to the tools they need to make decisions and give them a solid method for determining the right tool and technique for selection problems. This way, the CAD diagnosis rate will increase by determining the most effective method in these two interconnected areas. We present this article in accordance with the TRIPOD reporting checklist (available at https://cdt.amegroups.com/article/view/10.21037/cdt-24-237/rc).

Methods

Selected parameters/performance metrics of AI tools

Accuracy

The accuracy of a model is determined by dividing the total number of predictions by the number of total predictions. This ratio shows how well the model performs on a specific task. Higher percentages indicate greater overall performance when it comes to accuracy (34).

$$Accuracy = \left[\left(TP + TN \right) / \left(TN + TP + FN + FP \right) \right] \times 100\% \qquad [1]$$

True negative (TN): refers to situations in a binary classification where a model accurately predicts the negative category. True positive (TP): refers to situations in a binary

Table 1 Performance metrics of AI tools for CAD diagnosis

Performance metrics/AI tools	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	F1-score (%)	AUC-ROC
ResNet50	78.12 (20)	87 (21)	78.13 (20)	100 (20)	87.72 (20)	0.90 (22)
VGG16	84.38 (20)	93.33 (23)	83.33 (20)	100 (20)	90.91 (20)	0.91 (23)
DenseNet121	81.25 (20)	86.11 (23)	80.65 (20)	100 (20)	89.29 (20)	0.88 (23)
InceptionV3	84.38 (20)	79.25 (24)	100 (20)	80 (20)	88.89 (20)	0.93 (25)
SVM	91.5 (18)	95.0 (18)	97 (26)	87.6 (18)	92 (18,26)	0.856 (27)
KNN	91.9 (18)	99.8 (18)	87.7 (28)	83.9 (18)	85.75 (18,28)	0.880 (23)
RF	93.4 (18)	94.4 (18)	94.8 (28)	90.3 (18)	92.5 (18,28)	0.99 (18)
NB	59.3 (18)	77.9 (28)	70.9 (28)	86.7 (18)	78 (18,28)	0.889 (23)

AI, artificial intelligence; CAD, coronary artery disease; VGG16, visual geometry group; DenseNet121, dense convolutional network; SVM, support vector machine; KNN, k-nearest neighbors; RF, random forests; NB, naïve bayes; AUC-ROC, area under the receiver operating characteristic curve.

Table 2 Criteria of MPI techniques for CAD diagnosis

Criteria/MPI techniques	Specificity (%)	Sensitivity (%)	Radiation dose (mSv)	Cost of scan (€)	Study duration (minutes)
PET	86 (19)	85 (19)	9.7 (29)	1,192 (30)	30–40 (31)
SPECT	77 (19)	83 (19)	11.5 (29)	973 (30)	180–240 (31)
CMR	83 (19)	86 (19)	0 (32)	628 (33)	30–60 (32)

MPI, myocardial perfusion imaging; CAD, coronary artery disease; PET, positron emission tomography; SPECT, single photon emission computed tomography; CMR, cardiovascular magnetic resonance imaging.

classification where a model accurately predicts the positive category. False negative (*FN*): refers to a binary classification model that predicts the negative category wrongly. False positive (*FP*): refers to a binary classification model that predicts the positive category wrongly (34).

Specificity

Specificity is a metric used in binary classification to evaluate how well a model distinguishes the *TN* from the actual negatives (34).

$$Specificity = \left[TN/\left(TN+FP\right)\right] \times 100\% \quad [2]$$

Precision

By calculating the ratio of true positives to all predicted positives in a binary classification scenario, precision serves as a tool to assess the accuracy of positive predictions (34).

$$Precision = \left[TP/\left(TN+FP\right)\right] \times 100\% \quad [3]$$

Recall

In binary classification, recall, also referred to as sensitivity, is a parameter that assesses the model's efficacy in identifying positive instances. It is calculated as the ratio of true positives to the total of FN and TP (34).

$$Recall = \left[TP/\left(TN+FN\right)\right] \times 100\% \quad [4]$$

F1-score

In binary classification, the *F1*-score is a statistic that strikes a compromise between recall and precision. It evaluates a model's overall performance by taking the harmonic mean of these two metrics (34).

$$F1\text{-}score = \left[\left(2 * Recall * Precision\right)/\left(Recall + Precision\right)\right] \times 100\% \quad [5]$$

AUC-ROC

The area under the receiver operating characteristic curve, or AUC-ROC, is a metric used to assess how well a binary classification model performs. It produces a single value that expresses the model's ability to discriminate between positive and negative classes over a variety of thresholds (20).

Selected parameters/criteria of MPI techniques

Specificity

In MPI devices, specificity refers to the rate at which the test correctly identifies individuals without CAD, meaning that the test is less likely to produce a false-positive result.

Sensitivity

In MPI devices, sensitivity refers to the rate at which individuals with CAD are correctly identified. That is, it indicates the test's capacity to produce a true positive result, thus making the correct diagnosis without missing the disease.

Radiation dose

Nuclear imaging techniques (PET and SPECT) and CMR are important diagnostic methods in MPI. Radiation exposure occurs in the patient during nuclear imaging, which uses radioisotopes to monitor intrabody physiological processes. Exposure to radiation, even at low levels, can cause cellular DNA damage. CMR, on the other hand, works with magnetic fields and radio waves and does not contain ionizing radiation.

Cost of scan

The cost of scans defines the average processing cost of each MPI technique. Since radioisotopes and procedures used for nuclear imaging are generally expensive, CMR appears to be a more affordable method.

Study duration

For nuclear MPI, study duration includes the time it takes to inject the radioisotope, wait for it to distribute throughout the body, and then scan images. CMR scans are generally completed in less time. CMR scans that use contrast material may take longer than usual. In general, nuclear imaging may take longer, while CMR scans take less time.

Table 1 shows data on eight distinct AI tools (ResNet50, VGG16, DenseNet121, InceptionV3, SVM, KNN, RF, and NB) employed in MPI for CAD diagnosis. The data belongs to six important performance metrics (accuracy, precision, specificity, F1-score, recall, and AUC-ROC) for the analysis of AI tools. These data are obtained and referenced from open-access, peer-reviewed journals. The analyzed data presents the model's performance metrics, obtained from the testing data after dividing the dataset into training and test groups, approximately 80% and 20% of the total, respectively.

Table 2 shows data on three distinct MPI techniques (PET, SPECT, and CMR) for CAD diagnosis. The data belongs to five different criteria (specificity, sensitivity, radiation dose, cost of scan, and study duration) for the analysis of MPI techniques. These data were obtained from and referenced in open-access, peer-reviewed journal publications from recent years.

Study design

The data used in this study were collected based on the findings presented in open-access articles published in peer-reviewed journals and meticulously referenced. In obtaining the data for AI tools, the results of the same group of diseases who applied stress SPECT MPI and were diagnosed with CAD were taken into evaluation. For MPI techniques analyzed, data include the findings of the parameters/criteria for CAD diagnosis. This indicates that the data analyzed for each alternative belong to the same disease and diagnosis group. The alternative's performance metrics and criteria were compared, and CAD diagnosis rates and other MPI-based parameters were evaluated. A detailed analysis of each alternative was conducted in line with these performance metrics and criteria for ranking results using the fuzzy PROMETHEE method.

Figure 1 shows the steps of the methodology and design of the study. After determining the MPI techniques and AI tools used for CAD diagnosis, performance metrics and criteria that enable the analysis and comparison of alternatives were determined. Each parameter was weighted according to its importance level using the fuzzy linguistic scale, based on the experts' opinions.

Fuzzy logic & multi criteria decision making

Fuzzy logic is a mathematical notion that addresses uncertainty. It supports incomplete or uncertain data by accommodating degrees of truth, in contrast to conventional binary logic. Fuzzy logic, widely employed in AI and control systems, offers a flexible method of making decisions where exact definitions may be difficult (35).

Making decisions by taking into consideration a variety of factors is known as MCDM. It offers a methodical

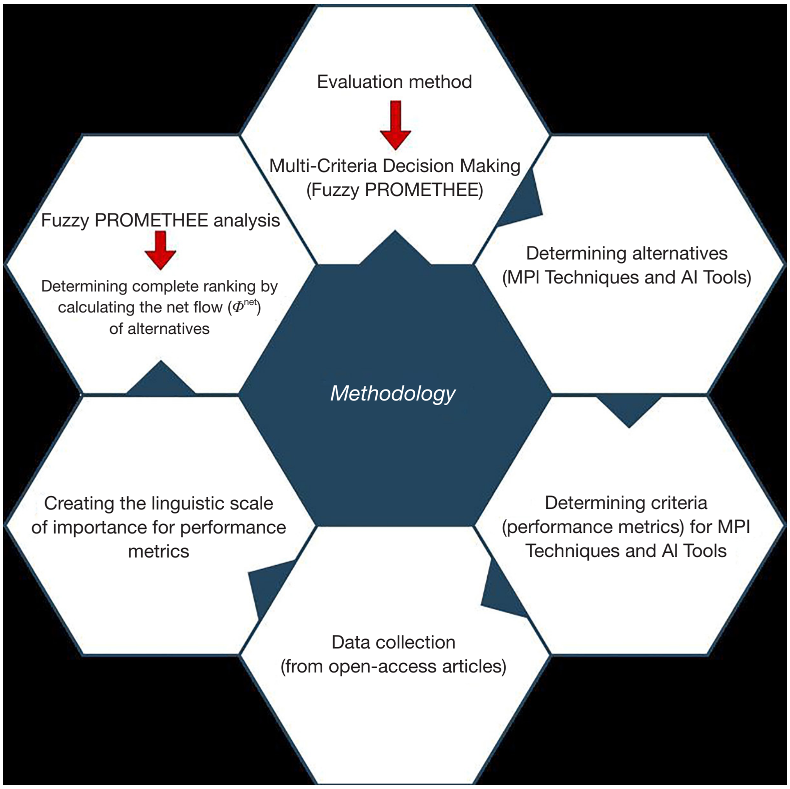

Figure 1 Flowchart of the study. PROMETHEE, preference ranking organization method for enrichment evaluation; MPI, myocardial perfusion imaging; AI, artificial intelligence.

framework for evaluating and ranking options according to how well they perform across a variety of criteria. It also facilitates a more thorough comparison and study of each analysis's criteria (36). The PROMETHEE technique differs from the other MCDM methods and can produce more sensitive results by providing different preference functions, such as linear function, V-shaped function, Gaussian function, level function, and U-shaped function, to the criteria for the comparison of the alternatives. The specifications of the weights of the alternatives have been determined in this work using a linguistic fuzzy scale. The linguistic fuzzy scale is used to consider the uncertainty and subjective assessments for decision in the determination of the weights of the criteria. The importance levels of the criteria are determined by linguistic expressions such as "very high", "high" and "moderate" instead of exact numbers, and these expressions are represented by fuzzy numbers. The importance levels of the criteria are determined according to the decision-maker's experiences, the impact of the criteria on the problem, and the priorities in the decision process. This approach allows for a more flexible and accurate assessment of uncertainties in the decision-making process. After converting these values to triangular fuzzy numbers, the data was defuzzified by applying the Yager index, which is a successfully used centroid method. Next, the net ranking result of the PROMETHEE approach was determined using the decision lab program by using the Gaussian preference function.

Steps of the PROMETHEE method

PROMETHEE is a decision-making process that evaluates and ranks possibilities according to several factors. Weighing the benefits and drawbacks of several options enables decision-makers to establish a preference structure and make well-informed decisions in challenging situations. It is developed by Brans *et al.* (37). There are 6 steps in PROMETHEE approach as below.

(I) Determine a specific preference function $p_j(d)$ for each alternative (AI tools or MPI techniques in this study) $j; a_j$.

(II) Determine the weight of each parameter (performance metrics and criteria) $w_t = (w_1, w_2, w_3, \ldots, w_k)$. If the decision maker determines that each parameter's weight is equally important, the parameters can all be specified equally.

(III) Determine the outranking relation π for all alternative pairs: $a_t, a_{t'} \in A$.

$$\pi(a_t, a_{t'}) = \sum_{k=1}^{k} w_k \left[p_k \left(f_k(a_t) - f_k(a_{t'}) \right) \right], \quad AXA \to [0,1] \qquad [6]$$

(IV) Determine the negative (entering) and positive (leaving) outranking flows Negative (entering) outranking flow for a_t:

$$\varphi^-(a_t) = \frac{1}{n-1} \sum_{t'=1, t \neq t'}^{n} \pi(a_{t'}, a_t) \qquad [7]$$

Positive (leaving) outranking flow for a_t:

$$\varphi^+(a_t) = \frac{1}{n-1} \sum_{t'=1, t \neq t'}^{n} \pi(a_t, a_{t'}) \qquad [8]$$

n denotes the number of alternatives, which are the AI tools and MPI techniques in our study. Each alternative is compared with $(n-1)$ number of another one. The positive (leaving) outranking flow $\varphi^+(a_t)$ refers to the strength of the alternatives $(a_t) \in A$ while the negative (entering) outranking flow $\varphi^+(a_t)$ refers to the weakness $(a_t) \in A$.

(V) Determine the partial pre order. In PROMETHEE I, alternatives a_t are preferred to $a_{t'}$ $(a_t P a_{t'})$ if they satisfies the one of the following conditions:

$$\begin{Bmatrix} \varphi^+(a_t) \succ \varphi^+(a_{t'}) \vee \varphi^-(a_t) \leq \varphi^-(a_{t'}) \\ \varphi^+(a_t) = \varphi^+(a_{t'}) \vee \varphi^-(a_t) \prec \varphi^-(a_{t'}) \end{Bmatrix} \qquad [9]$$

If there are two alternatives a_t and $a_{t'}$ with similar or equal positive (leaving) and negative (entering) flows, a_t is indifferent to $a_{t'}$ $(a_t I a_{t'})$.

$$\varphi^+(a_t) = \varphi^+(a_{t'}) \vee \varphi^-(a_t) = \varphi^-(a_{t'}), \quad \text{if } (a_t I a_{t'}) \qquad [10]$$

a_t is incomparable to $a_{t'}$ $(a_t R a_{t'})$ if;

$$\begin{Bmatrix} \varphi^+(a_t) \succ \varphi^+(a_{t'}) \vee \varphi^-(a_t) \succ \varphi^-(a_{t'}) \\ \varphi^+(a_t) \prec \varphi^+(a_{t'}) \vee \varphi^-(a_t) \prec \varphi^-(a_{t'}) \end{Bmatrix} \qquad [11]$$

(VI) Determine the net outranking flow for each of the alternatives using the Eq. [12].

$$\varphi^{net}(a_t) = \varphi^+(a_t) - \varphi^-(a_t) \qquad [12]$$

It can obtain the entire pre-order through the net flow and utilization of PROMETHEE II.

$$\varphi^{net}(a_t) \succ \varphi^{net}(a_{t'}), \quad \text{if } (a_t P a_{t'}) \qquad [13]$$

$$\varphi^{net}(a_t) = \varphi^{net}(a_{t'}), \quad \text{if } (a_t I a_{t'}) \qquad [14]$$

As a result, the most effective alternative is the one having the higher $\varphi^{net}(a_t)$ (the net flow) value.

In this study, 5-scale triangular linguistic fuzzy sets are used to determine the weights for MPI techniques and performance metrics of the AI tools: Very High/(0.75, 1, 1), High/(0.5, 0.75, 1), Moderate/(0.25, 0.5, 0.75), Low/(0, 0.25, 0.5), and Very Low/(0, 0, 0.25). The weights of all performance metrics of the AI tools (accuracy, specificity, precision, F1-score, recall, and AUC-ROC) are assigned as 'Very High' because they are all the main metrics for analyzing AI tools in SPECT MPI. However, in the evaluation of MPI techniques, the weights of sensitivity and specificity were assigned as 'Very High', while the weights of radiation dose, cost of scan, and study duration were assigned as 'High'. This decision was made considering that among the criteria, those that would have a greater impact on the CAD diagnosis rate were more important. The opinions of the cardiologist were obtained to determine the linguistic scale of importance weights of each performance metric and criterion.

The fuzzy scale of weights for the criteria of the MPI techniques and performance metrics of the AI tools has been consolidated into a single point using the Yager index due to its ability to compare fuzzy values by decision-makers rationally. The preference function utilized in this study is the Gaussian function applied to each criterion. This choice is grounded in its capacity to offer a continuous probability distribution that exhibits symmetry around its mean.

Results

Each alternative (AI tools and MPI techniques for CAD diagnosis) was numerically compared based on performance metrics and/or selected criteria, and as a result of these comparisons, positive and negative outrankıng flow values are calculated and the net flow (φ^{net}) of each alternative was determined. The positive outranking flow represents the strengths of the alternatives, while the negative outranking flow reflects their weaknesses. The net flow, therefore, yields the results of net ranking; the higher the net flow, the more effective the alternative(s). The importance levels of performance metrics are taken into consideration in these

Table 3 Complete ranking of AI tools

Complete ranking	AI tools	Positive outranking flow	Negative outranking flow	Net flow
1	RF	0.4926	0.1147	0.3778
2	SVM	0.4548	0.1108	0.3440
3	VGG16	0.3759	0.1867	0.1892
4	KNN	0.3921	0.2745	0.1176
5	DenseNet121	0.2678	0.3086	−0.0408
6	ResNet50	0.2411	0.3826	−0.1415
7	InceptionV3	0.2397	0.3896	−0.1499
8	NB	0.0303	0.7266	−0.6963

AI, artificial intelligence; RF, random forests; SVM, support vector machine; VGG16, visual geometry group; KNN, k-nearest neighbors; DenseNet121, dense convolutional network; ResNet50, residual neural network; NB, naïve bayes.

calculations.

Results of AI tools in MPI for CAD diagnosis

The ranking outcomes for the AI tools utilized in MPI for the diagnosis of CAD are detailed in *Table 3*. The findings derived from the fuzzy PROMETHEE analysis reveal that the RF stands out as the most optimal option as an AI tool for MPI in the diagnosis of CAD disease with a net flow of 0.3778. This preference is substantiated by noteworthy performance metrics, including accuracy (93.4%), specificity (94.4%), precision (94.8%), F1-score (92.5%), and AUC-ROC (0.99). However, recall (90.3%) remained below the average among performance metrics. Following RF, SVM emerges as the second optimal choice with a net flow of 0.3440, showcasing results in accuracy (91.5%), specificity (95.0%), precision (97.0%), F1-score (92.0%), recall (87.6%), and AUC-ROC (0.856). Conversely, NB lags in all performance metrics, encompassing accuracy (59.3%), specificity (77.9%), precision (70.9%), recall (86.7%), F1-score (78.0%), and AUC-ROC (0.889), falling below the average with a net flow of −0.6963.

Based on the selected parameters, the preference ranking of the AI tools for CAD diagnosis was determined as RF, SVM, VGG16, KNN, DenseNet121, ResNet50, InceptionV3, and NB, respectively.

The rainbow diagram obtained from PROMETHEE indicates the advantages and disadvantages of each alternative and ranks alternatives from the most effective to the least effective. *Figure 2* depicts a detailed rainbow ranking of AI tools and their associated performance metrics, elucidating the factors that render AI tools superior or weak for MPI in diagnosing CAD. The performance metrics above the AI tools (0 thresholds) on the graph highlight their superiority, while those below delineate their weaknesses. This diagram shows that RF was determined as the most effective AI tool and only recall emerged as a weakness in the performance metrics.

Results of MPI techniques for CAD diagnosis

In the second analysis, the findings derived from the fuzzy PROMETHEE analysis reveal that CMR stands out as the most optimal technique for MPI in the diagnosis of CAD with a net flow of 0.3666. The ranking outcomes for MPI techniques for the diagnosis of CAD are detailed in *Table 4*. This preference is substantiated by noteworthy criteria, including sensitivity (86.0%), radiation dose (0 mSv), and cost of scan (€628). However, some criteria such as specificity (83.0%) and study duration (30-60 minutes) fall below the PET. Following CMR, PET scan emerges as the second optimal choice with a net flow of −0.0764, showcasing results in specificity (86.0%), sensitivity (85.0%), radiation dose (9.7 mSv), cost of scan (€1192), and study duration (30–40 minutes). Conversely, SPECT lags in all criteria, encompassing specificity (77.0%), sensitivity (83.0%), radiation dose (11.5 mSv), cost of scan (€973) and study duration (180–240 minutes), falling below the average with a net flow of −0.2902.

Table 4 shows the complete ranking of the MPI techniques to diagnose CAD. Based on the selected parameters, the preference ranking was determined as CMR, PET, and SPECT, respectively.

Figure 3 depicts a detailed rainbow ranking of MPI

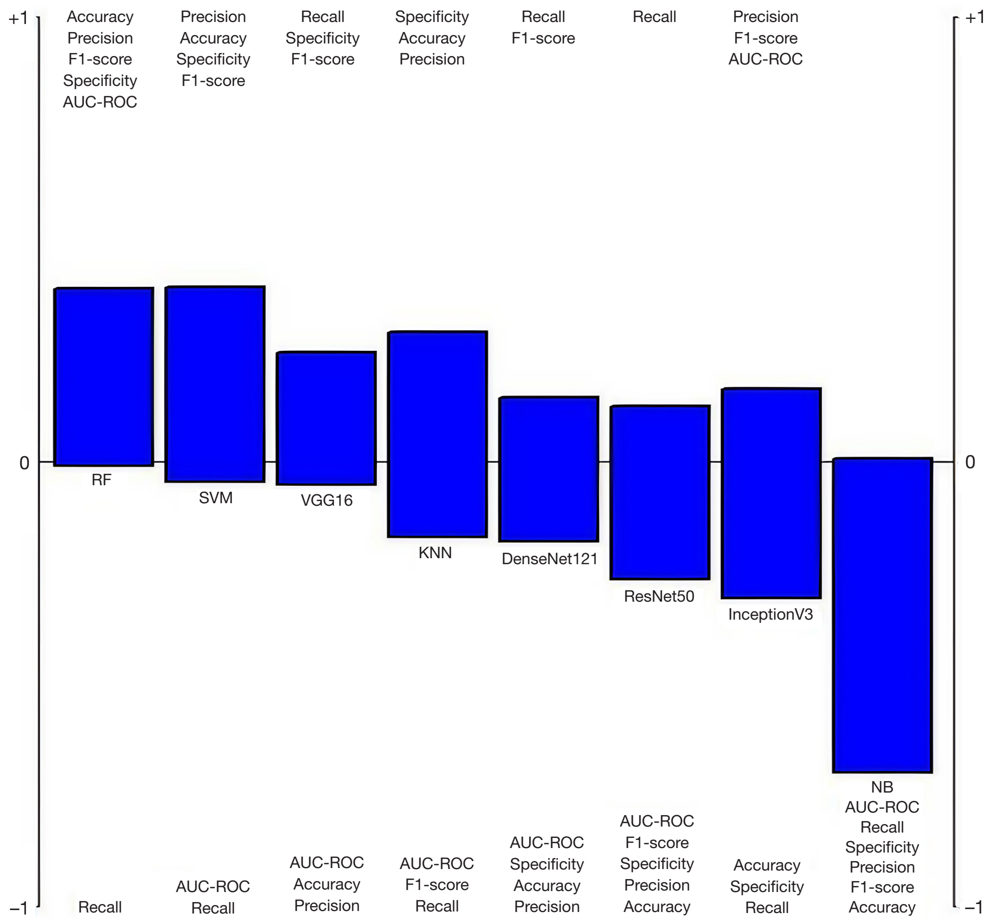

Figure 2 PROMETHEE rainbow diagram for ranking of AI tools. RF, random forests; SVM, support vector machine; VGG16, visual geometry group; KNN, k-nearest neighbors; DenseNet121, dense convolutional network; ResNet50, residual neural network; NB, naïve bayes; AUC-ROC, area under the receiver operating characteristic curve; PROMETHEE, preference ranking organization method for enrichment evaluation.

Table 4 Complete ranking of MPI techniques

Complete ranking	MPI techniques	Positive outranking flow	Negative outranking flow	Net flow
1	CMR	0.4579	0.0913	0.3666
2	PET	0.1982	0.2746	–0.0764
3	SPECT	0.0917	0.3819	–0.2902

MPI, myocardial perfusion imaging; CMR, cardiovascular magnetic resonance imaging; PET, positron emission tomography; SPECT, single photon emission computed tomography.

techniques and their associated criteria, elucidating the factors that render techniques superior or weak for MPI in diagnosing CAD. The criteria positioned above the MPI techniques on the graph highlight their superiority, while those below delineate their weaknesses. This evaluation results in a ranking of MPI techniques from the most effective to the least effective. This diagram shows that CMR is the most effective MPI technique, showing

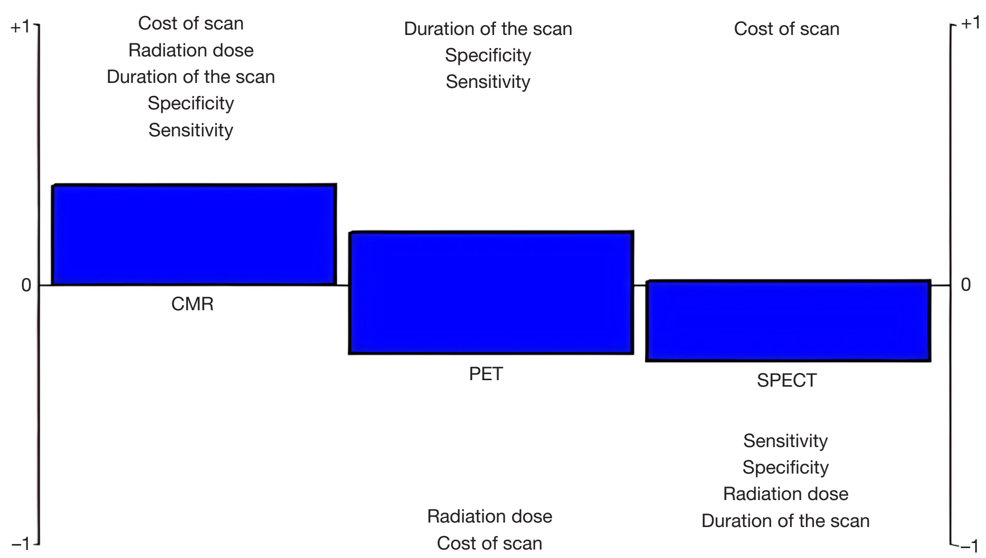

Figure 3 PROMETHEE rainbow diagram for ranking of MPI techniques. CMR, cardiovascular magnetic resonance imaging; PET, positron emission tomography; SPECT, single photon emission computed tomography; PROMETHEE, preference ranking organization method for enrichment evaluation; MPI, myocardial perfusion imaging.

superiority in all criteria.

Discussion

This study conducts a thorough comparison of AI tools used in MPI for diagnosing CAD, along with a detailed comparison and analysis of MPI techniques.

While existing literature has individually compared the results of ML and CNN tools, there is currently no comprehensive study beyond this research that encompasses eight distinct AI tools and assesses performance metrics for accurate diagnosis. This extensive investigation incorporates parameters such as accuracy, specificity, precision, F1-score, recall, and AUC-ROC through multi-criteria decision-making. The identification of the most effective AI tool will contribute to enhancing the diagnostic rate for CAD. Consequently, the results of the study are both in line with the literature and supportive. The findings are derived from the analysis of images obtained through SPECT-MPI.

According to the results of MPI techniques, which is another analysis of the study, CMR emerges as a good alternative to nuclear imaging techniques. For patients who do not want to be exposed to radiation or when continuous imaging is required, it can be preferred as an option that helps to prevent the effects of radioactive substances from nuclear imaging techniques. It also offers a cheaper scan fee. However, CMR has limitations for patients who are allergic to contrast agents, have kidney problems, or have metallic implants. When the radiation dose and scan fee are not included in the analysis, PET stands out with its even better specificity value. For this reason, due to its appeal to a wider range of patients, it is important to develop nuclear imaging techniques.

The study found that, despite being the most frequently preferred nuclear MPI method in CAD evaluation, SPECT was the least effective method. This analytical approach, which allows the analysis of alternatives according to multi-criteria and different importance weights, has reached these findings. The frequent preference for SPECT in CAD evaluation can be explained by various factors such as technological differences, clinical practice and accessibility, cost, and ease of use.

Other MPI techniques not examined in the study are known to be important for CAD analysis, and in some cases, using multiple methods in diagnostic processes is beneficial. However, this study highlights that AI can produce results with accuracy that can reduce the need for multiple MPI techniques. Although reliable systems that fully meet this need do not yet exist, this study can provide important guidance for the development of such systems. In light of recent advancements in nuclear imaging devices and emerging AI tools, regular updates and reevaluation of data may be essential. The study acknowledges some limitations. The performance metrics of the high-importance AI tools analyzed in the study and the findings regarding the criteria for MPI techniques are limited due to the restricted data sources and the fact that they are rarely used in the same population. Nevertheless, the findings will serve as a valuable guide for identifying and addressing the limitations of current AI tools and nuclear MPI techniques.

Conclusions

The findings of this study reveal that the RF algorithm exhibits superior performance among the AI tools applied using SPECT MPI in the diagnosis of CAD, as indicated by the high net flow value in the fuzzy PROMETHEE results. It was also determined that CMR imaging yielded better results compared to nuclear MPI techniques. This research proposes a novel approach to determining the most effective AI tools and MPI techniques in CAD diagnosis by considering additional criteria and importance weights beyond the current applications. Consequently, this study underscores the pivotal role of MPI in diagnosing CAD and highlights the transformative impact of AI in enhancing the diagnostic accuracy and efficiency of imaging techniques. This study focuses on nuclear MPI techniques and compares them only with one non-nuclear alternative CMR. To generalize the findings more comprehensively, it is recommended that future studies include different diagnostic techniques in the analysis. Such extended analysis will be useful to evaluate the performance of AI tools in CAD diagnosis and obtain more comprehensive results.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://cdt.amegroups.com/article/view/10.21037/cdt-24-237/rc

Peer Review File: Available at https://cdt.amegroups.com/article/view/10.21037/cdt-24-237/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://cdt.amegroups.com/article/view/10.21037/cdt-24-237/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Celermajer DS, Chow CK, Marijon E, et al. Cardiovascular disease in the developing world: prevalences, patterns, and the potential of early disease detection. J Am Coll Cardiol 2012;60:1207-16.
2. Pazoki R, Dehghan A, Evangelou E, et al. Genetic Predisposition to High Blood Pressure and Lifestyle Factors: Associations With Midlife Blood Pressure Levels and Cardiovascular Events. Circulation 2018;137:653-61.
3. McAloon CJ, Boylan LM, Hamborg T, et al. The changing face of cardiovascular disease 2000-2012: An analysis of the world health organisation global health estimates data. Int J Cardiol 2016;224:256-64.
4. Crystal GJ, Assaad SI, Heerdt PM. Cardiovascular physiology: integrative function. In: Hemmings HC, Egan TD, editors. Pharmacology and Physiology for Anesthesia. 2019, pp. 473-519. Elsevier.
5. Hussain S, Mubeen I, Ullah N, et al. Modern Diagnostic Imaging Technique Applications and Risk Factors in the Medical Field: A Review. Biomed Res Int 2022;2022:5164970.
6. Klein R, Celiker-Guler E, Rotstein BH, et al. PET and SPECT Tracers for Myocardial Perfusion Imaging. Semin Nucl Med 2020;50:208-18.
7. Emrich T, Halfmann M, Schoepf UJ, et al. CMR for myocardial characterization in ischemic heart disease: state-of-the-art and future developments. Eur Radiol Exp 2021;5:14.
8. Symons R, Zimmerman SL, Bluemke DA. CMR and CT of the Patient With Cardiac Devices: Safety, Efficacy,

and Optimization Strategies. JACC Cardiovasc Imaging 2019;12:890-903.
9. Prince JL, Links JM. Medical imaging signals and systems. 2nd ed. Pearson; 2023.
10. Dobrucki LW, Sinusas AJ. PET and SPECT in cardiovascular molecular imaging. Nat Rev Cardiol 2010;7:38-47.
11. Nudi F, Nudi A, Biondi-Zoccai G, et al. Bridging the Atlantic gap in clinical guidelines for non-ST-elevation acute coronary syndromes. J Nucl Cardiol 2018;25:780-4.
12. Papandrianos NI, Apostolopoulos ID, Feleki A, et al. AI-based classification algorithms in SPECT myocardial perfusion imaging for cardiovascular diagnosis: a review. Nucl Med Commun 2023;44:1-11.
13. Juarez-Orozco LE, Klén R, Niemi M, et al. Artificial Intelligence to Improve Risk Prediction with Nuclear Cardiac Studies. Curr Cardiol Rep 2022;24:307-16.
14. Dey D, Slomka PJ, Leeson P, et al. Artificial Intelligence in Cardiovascular Imaging: JACC State-of-the-Art Review. J Am Coll Cardiol 2019;73:1317-35.
15. Papandrianos N, Papageorgiou E. Automatic diagnosis of coronary artery disease in SPECT myocardial perfusion imaging employing deep learning. Appl Sci 2021;11:6362.
16. Chartrand G, Cheng PM, Vorontsov E, et al. Deep Learning: A Primer for Radiologists. Radiographics 2017;37:2113-31.
17. Erickson BJ, Korfiatis P, Akkus Z, et al. Machine Learning for Medical Imaging. Radiographics 2017;37:505-15.
18. Cantoni V, Green R, Ricciardi C, et al. Comparing the Prognostic Value of Stress Myocardial Perfusion Imaging by Conventional and Cadmium-Zinc Telluride Single-Photon Emission Computed Tomography through a Machine Learning Approach. Comput Math Methods Med 2021;2021:5288844.
19. Xu J, Cai F, Geng C, et al. Diagnostic Performance of CMR, SPECT, and PET Imaging for the Identification of Coronary Artery Disease: A Meta-Analysis. Front Cardiovasc Med 2021;8:621389.
20. Magboo VPC, Magboo MSA. Diagnosis of Coronary Artery Disease from Myocardial Perfusion Imaging Using Convolutional Neural Networks. Procedia Computer Science 2023;218:810-7.
21. Táborský M, Richter D, Tonar Z, et al. Evaluation of later morphologic alterations in renal artery wall and renal nerves in response to catheter-based renal denervation in sheep: comparison of the single-point and multiple-point ablation catheters. Physiol Res 2018;67:891-901.
22. Juarez-Orozco LE, Martinez-Manzanera O, van der Zant FM, et al. Deep Learning in Quantitative PET Myocardial Perfusion Imaging: A Study on Cardiovascular Event Prediction. JACC Cardiovasc Imaging 2020;13:180-2.
23. Papandrianos NI, Feleki A, Papageorgiou EI, et al. Deep Learning-Based Automated Diagnosis for Coronary Artery Disease Using SPECT-MPI Images. J Clin Med 2022;11:3918.
24. Papandrianos NI, Feleki A, Moustakidis S, et al. An explainable classification method of SPECT myocardial perfusion images in nuclear cardiology using deep learning and grad-CAM. Appl Sci 2022;12:7592.
25. AlOthman AF, Sait ARW, Alhussain TA. Detecting Coronary Artery Disease from Computed Tomography Images Using a Deep Learning Technique. Diagnostics (Basel) 2022;12:2073.
26. Prabha AJ, Bhargavi R, Harish B. An efficient machine learning model for prediction of dyslexia from eye fixation events. New Approaches Eng Res 2021;10:171-9.
27. Li J, Yang G, Zhang L. Artificial Intelligence Empowered Nuclear Medicine and Molecular Imaging in Cardiology: A State-of-the-Art Review. Phenomics 2023;3:586-96.
28. Ricciardi C, Cantoni V, Improta G, et al. Application of data mining in a cohort of Italian subjects undergoing myocardial perfusion imaging at an academic medical center. Comput Methods Programs Biomed 2020;189:105343.
29. Al Badarin FJ, Spertus JA, Bateman TM, et al. Drivers of radiation dose reduction with myocardial perfusion imaging: A large health system experience. J Nucl Cardiol 2020;27:785-94.
30. Mimouni M, Bulsei J, Darlington M, et al. Cost-effectiveness of 82-Rubidium PET myocardial perfusion imaging for the diagnosis of myocardial ischemia depending on the prevalence of coronary artery disease. EJNMMI Res 2023;13:9.
31. El-Tallawi KC, Aljizeeri A, Nabi F, et al. Myocardial Perfusion Imaging Using Positron Emission Tomography. Methodist Debakey Cardiovasc J 2020;16:114-21.
32. Arai AE, Schulz-Menger J, Shah DJ, et al. Stress Perfusion Cardiac Magnetic Resonance vs SPECT Imaging for Detection of Coronary Artery Disease. J Am Coll Cardiol 2023;82:1828-38.
33. Patel AR, Salerno M, Kwong RY, et al. Stress Cardiac Magnetic Resonance Myocardial Perfusion Imaging: JACC Review Topic of the Week. J Am Coll Cardiol 2021;78:1655-68.
34. Vujović Ž. Classification model evaluation metrics. International Journal of Advanced Computer Science and

Applications 2021;12:599-606.
35. Zadeh LA, Yuan B, Klir GJ. Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A. Zadeh. NJ: World Scientific Publishing 1996:826.
36. Erdagli H, Ozsahin DU, Uzun B. Evaluation and simulation of breast cancer imaging devices using multi-criteria decision theory. Journal of Instrumentation 2020;15:C05029.
37. Brans JP, Vincke P, Mareschal B. How to select and how to rank projects: The PROMETHEE method. European Journal of Operational Research 1986; 24:228-38.

Cite this article as: Erdagli H, Uzun Ozsahin D, Uzun B. Evaluation of myocardial perfusion imaging techniques and artificial intelligence (AI) tools in coronary artery disease (CAD) diagnosis through multi-criteria decision-making method. Cardiovasc Diagn Ther 2024;14(6):1134-1147. doi: 10.21037/cdt-24-237