

A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system

Hamid Reza Marateb, Sobhan Goudarzi

Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran

Background: Coronary heart diseases/coronary artery diseases (CHDs/CAD), the most common form of cardiovascular disease (CVD), are a major cause for death and disability in developing/developed countries. CAD risk factors could be detected by physicians to prevent the CAD occurrence in the near future. Invasive coronary angiography, a current diagnosis method, is costly and associated with morbidity and mortality in CAD patients. The aim of this study was to design a computer-based noninvasive CAD diagnosis system with clinically interpretable rules. **Materials and Methods:** In this study, the Cleveland CAD dataset from the University of California UCI (Irvine) was used. The interval-scale variables were discretized, with cut points taken from the literature. A fuzzy rule-based system was then formulated based on a neuro-fuzzy classifier (NFC) whose learning procedure was speeded up by the scaled conjugate gradient algorithm. Two feature selection (FS) methods, multiple logistic regression (MLR) and sequential FS, were used to reduce the required attributes. The performance of the NFC (without/with FS) was then assessed in a hold-out validation framework. Further cross-validation was performed on the best classifier. **Results:** In this dataset, 16 complete attributes along with the binary CHD diagnosis (gold standard) for 272 subjects (68% male) were analyzed. MLR + NFC showed the best performance. Its overall sensitivity, specificity, accuracy, type I error (α) and statistical power were 79%, 89%, 84%, 0.1 and 79%, respectively. The selected features were "age and ST/heart rate slope categories," "exercise-induced angina status," fluoroscopy, and thallium-201 stress scintigraphy results. **Conclusion:** The proposed method showed "substantial agreement" with the gold standard. This algorithm is thus, a promising tool for screening CAD patients.

Key words: Classification, clinical prediction rule, coronary artery disease, data mining, fuzzy logic

How to cite this article: Marateb HR, Goudarzi S. A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system. *J Res Med Sci* 2015;20:214-23.

INTRODUCTION

Coronary heart disease also known as coronary artery disease (CAD) is a chronic disease in which the coronary arteries, responsible for transporting oxygenated blood to heart muscles, get narrowed and are not able to convey enough fresh blood to this blood-pumping organ.^[1,2] Narrowing of blood vessels is usually due to arteriosclerosis, a common arterial disease in which increased areas of degeneration and cholesterol (CHOL) deposit plaques form on the inner surfaces of the arteries blocking blood flow.^[3,4] In case of reduced blood supply, the heart does not receive enough oxygen and nutrition to operate properly resulting in angina pectoris and heart attack. The symptomatic or asymptomatic reduction in coronary artery flow, may occur with exercise or at rest, and may end up with a myocardial infarction, depending on the severity of obstruction and the speed of its development.^[5,6] CAD, the most common form of cardiovascular disease (CVD), has the prevalence of 6.9% in men and 6% in women.^[7]

Despite the background of increasing health care expenses, CAD has a significant influence on global economics as a principal cause of disability and loss of efficiency.^[8] CAD is a major cause of death and disability in developed countries. Although CAD mortality rates have dropped over the past four decades, CAD remains responsible for about one-third of all deaths in people over age 35.^[6,7,9]

According to the World Health Organization, more than 60% of the global burden of the CAD occurs in developing countries.^[10] CAD is the leading cause of death worldwide: 3.8 million men and 3.4 million women each year.^[11] Overall, the prevalence of CAD in Iran (≥ 20 years old) was reported within 1.1-36.0% in different studies.^[12] The age-adjusted (≥ 30 years old) prevalence of CAD was reported as 21.8% (22.3% in women and 18.8% in men) in Tehran, Iran.^[13] The prevalence of CAD among people aged 35-79 years was reported to be 19.4% (21.9% in women and 16.0% in men) in an urban population in

Address for correspondence: Dr. Hamid Reza Marateb, Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Hezar Jerib St., Isfahan 81744, Isfahan, Iran. E-mail: h.marateb@eng.ui.ac.ir

Received: 08-09-2014; **Revised:** 28-10-2014; **Accepted:** 27-01-2015

Isfahan, Iran.^[14] Over the past decade, new scientific evidence strongly supporting the role of preventive interventions in the maintenance of health has focused much needed attention and efforts on cardiovascular prevention.^[15] Thus, the prevention of CAD is a major goal and has been focused on many international health programs.

Coronary heart disease risk assessment could be detected by the physicians to predict the CAD occurrence in the near future to control its progress in patients.^[16] One of the methods of CAD risk assessment is the investigation of its risk factors.^[4,17] Hypertension, high levels of low-density lipoprotein cholesterol (LDL-C), low levels of high-density lipoprotein cholesterol (HDL-C), high total cholesterol, high triglycerides, diabetes mellitus, smoking, obesity, aging, gender, physical inactivity, age, socioeconomic and psychological stress, family history of CAD, and various genetic factors^[8,18-20] are some of the CAD risk factors reported in the literature. A variety of meta-analysis studies have been performed on other CAD risk factors. For example, it was shown that depression,^[21] job strain,^[22] anxiety,^[23] a diet poor in fruit and vegetable^[24] and frequent fried-food consumption^[25] significantly increase the risk of CAD.

Medical diagnosis is a difficult and complex task, and computer-aided diagnosis systems are technologies designed to decrease observational oversight.^[26] A computer-aided diagnosis would be desirable for performing the classification and decision procedures since the computers can store large amount of data without distortion over long periods of time while performing complex logical and mathematical operations at very high speed.^[27,28] Computer-based medical diagnosis systems have been promoted for their potential to improve the quality of health care, including their application to support and improve clinical decisions.^[29,30]

Computer-aided medical diagnosis is usually performed via classification, also known as “supervised learning.” In machine learning, “classification” is the problem of identifying to which category (class) a new observation belongs, based on a training set containing observations (features) whose category membership is known (gold standard).^[31] A large number of classifiers have been proposed in the literature. Most of them use a black-box modeling approach without paying attention to the underlying mechanisms.^[32] Examples of these classifiers are linear and quadratic discriminant analysis, support vector machines, *k*-nearest neighbors, Naïve Bayesian classifier. Accordingly, these classifiers cannot be clinically interpreted which is not desirable in clinical applications.^[33] Rule-based classification systems such as the decision tree and its variants, on the other hand, can provide interpretable classification rules.^[34,35]

The medical knowledge and the resulting diagnosis are pervaded by uncertainty. Fuzzy set theory on the other hand was conceived with the formalization of the vague knowledge.^[36,37] It is very difficult to define sharp borders between various symptoms and diseases. Thus, the framework of the fuzzy system is very useful to deal with the absence of sharp boundaries of the sets of symptoms, diagnosis, and phenomena of diseases.^[38,39] Also, fuzzy logic is a useful tool for building expert systems for decision-making in the field of medical diagnosis.^[40] Accordingly, fuzzy rule-based classification systems are now quite popular in the field of medical diagnosis.^[41-43] These systems create clinically interpretable rules that take into account the overlap between different diagnosis classes. Therefore, these systems are proved to have better performance in comparison with that of crisp rule-based systems (such as decision tree).^[44,45]

In this paper, a fuzzy rule-based system was designed to diagnose CAD based on a limited set of features recorded noninvasively in a case-control study. The rest of the paper is organized as follow: Information about the study population, recorded features, and the data mining methodology is given in the next section. Then, the results of the classification and the extracted rules are reported in the “results” section and finally, the discussion about the clinical validity of the proposed system and the comparison with other classifiers will be provided in the “discussion” section.

MATERIALS AND METHODS

Experimental methods

In this work, the CAD dataset from the University of California (UCI, Irvine), which is available online (<https://archive.ics.uci.edu/ml/datasets/Heart+Disease>), taken from the Cleveland Clinic Foundation datasets, is used for training and testing the proposed system.^[46-49] This dataset has been considered as a benchmark for various computer-aided CAD diagnosis systems.^[50] This database consisted of 303 records with 76 attributes (features).

The experimental protocol can be found elsewhere in details.^[47,51] However, it is briefly mentioned as follows. A number of 303 consecutive patients referred for coronary angiography at the Cleveland Clinic between May 1981 and September 1984 participated in the experiment. No patient had a history or electrocardiographic evidence of prior myocardial infarction or known valvular or cardiomyopathic diseases. The following attributes were collected: Age, gender, resting blood pressure (restbps), CHOL, fasting blood sugar (FBS), resting electrocardiographic results (restecg), active smoker type, number of cigarettes per day, number of years as a smoker, family history of CAD (famhist), chest pain type (cp). In addition, patients underwent 3 noninvasive

tests as part of a research protocol. The results of these tests (exercise electrocardiogram, thallium scintigraphy and cardiac fluoroscopy) were not interpreted until after the invasive gold standard test (coronary angiograms) had been read. The CAD status was considered when narrowing of at least one of the coronary arteries was more than 50%, shown by angiography.^[52] The definition of attributes, their measurement scales and categories in the raw dataset were listed in Table 1.

Preprocessing

In our study, discretization was used on the attribute intervals. In machine learning, discretization refers to the process of converting or partitioning continuous attributes to discretized variables. Discretization is typically used as a preprocessing step for machine learning algorithms because it can significantly impact the performance of classification algorithms used for data analysis.^[53,54] Discretization is a technique to find cut points to partition the range into a small number of intervals by maintaining

good class consistency.^[55-57] We used a set of cut points taken from the literature, indicating the diagnosis properties of individual attributes as below to create ordinal-scale variables:

1. "age:" "Young adult" (18-35 years old), "middle-aged adults" (36-55 years old), and "older adults" (older than 55 years old);^[58]
2. "CHOL:" "Desirable" (<200 mg/dL), "borderline high" (200-239 mg/dL) and "high" (\geq 240 mg/dL);^[17]
3. "Trestbps:" "Low" (90-100 mmHg), "normal" (100-120 mmHg), "prehypertension" (120-139 mmHg), "stage 1 hypertension" (140-159 mmHg) and "stage 2 hypertension" (\geq 60 mmHg);^[59]
4. "Cigs:" Five categories 0, 1-9, 10-19, 20-39, and \geq 40;^[60]
5. "Thalrest:" The following categories were extracted ("excl't," "good," "above av," "average," "below av" and "poor") based on resting heart rate (HR) chart considering gender, age and resting HR;^[61]
6. "Oldpeak" and "thalach:" The ST depression was divided by the maximum exercise HR (ST/HR slope),

Table 1: The attributes of the raw Cleveland CAD dataset

Attribute	Measurement scale	Definition	Categories*
Age	Interval	Age in years	-
Gender	Nominal	Sex	Male/female
Trestbps	Interval	Resting blood pressure (mmHg)	-
CHOL	Interval	Serum CHOL (mg/dL)	-
FBS	Nominal	FBS >120 (mg/dL)	True/false
Restecg	Nominal	Resting electrocardiographic results	(1) Normal; (2) having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of >0.05 mV); (3) probable or definite left ventricular hypertrophy by Estes' criteria
Thalrest	Interval	Resting heart rate (bpm)	-
Smoke	Nominal	Active smoker type	Yes/no
Cigs	Interval	Number of cigarettes per day	-
years	Interval	Number of years as a smoker	-
Famhist	Nominal	Family history of CAD	Yes/no
Cp**	Nominal	Chest pain type	(1) Typical angina pectoris; (2) atypical angina; (3) nonanginal pain; (4) no pain
Tpeakbps	Interval	Peak exercise systolic blood pressure (mmHg)	-
Tpeakbpd	Interval	Peak exercise diastolic blood pressure (mmHg)	-
Thalach	Interval	Maximum exercise heart rate achieved (bpm)	-
Exang	Nominal	Exercise-induced angina	Yes/no
Oldpeak	Interval	ST depression induced by exercise relative to rest	-
Slope	Ordinal	The slope of the peak exercise ST segment	(1) Upsloping; (2) flat; (3) downsloping
Ca	Interval	Number of major vessels (0-3) colored by fluoroscopy	-
Thal***	Nominal	Thallium-201 stress scintigraphy	(3) Normal; (6) fixed defect; (7) reversible defect
Num	Nominal	Diagnosis of heart disease (angiographic disease status)	(1) Normal: <50% diameter narrowing; (2) CAD >50% diameter narrowing

*The categories were shown for nominal or ordinal features; **(1) Typical angina pectoris: Pain that occurs in the anterior thorax, neck, shoulders, jaw, or arms is precipitated by exertion and relieved within 20 min by rest. (2) Atypical angina: Pain in one of the above locations and either not precipitated by exertion or not relieved by rest within 20 min. (3) Nonanginal pain: Pain not located in any of the above locations, or if so located not related to exertion, and lasting less than 10 s or longer than 30 min. (4) No pain; ***(1) Normal, (2) Fixed abnormality (defects observed during exercise that persisted at redistribution), and (3) Reversible abnormality (defects present during exercise and significantly corrected during redistribution). CAD = Coronary artery disease; CHOL = Cholesterol; FBS = Fasting blood sugar

also known as HR adjustment of exercise-induced ST segment depression.^[62,63] Then, the cut points of 0.01 and 0.02 $\mu\text{V}/\text{beat}/\text{min}$ were used to derive three ordinal classes: “low,” “medium” and “high.”

Fuzzy rule-based system

Rule generation includes both rule extraction and refinement. One of the methods for extracting rules from the input-out data is neuro-fuzzy rule generation.^[64,65] It combines the powerful description of fuzzy classification techniques with the learning capabilities of Neural Networks. In the neuro-fuzzy classification methods, the feature space is partitioned into multiple fuzzy subspaces that are controlled by fuzzy if-then rules. For determining an optimum fuzzy region, the parameters of the fuzzy rules should be optimized.^[66] The classifier consists of the following layers: Fuzzy membership, fuzzification, defuzzification, normalization and output.^[67] The K-means clustering method was used to obtain the initial parameters and to formulate the fuzzy if-then rules.^[65]

A fuzzy classification rule R_i , which describes the relation between the input feature space and the classes, can be defined as follows:

R_i : If x_{p_1} is ϕ_{i_1} and ... and x_{p_j} is ϕ_{i_j} and R and x_{p_m} is ϕ_{i_m} then class is out_k .

where x_{pj} is the j^{th} feature of the p^{th} sample, out_k is the k^{th} class label, and ϕ_{ij} is the fuzzy set of the j^{th} feature in the i^{th} rule.^[67,68] In the neuro-fuzzy method used in this study, scaled conjugate gradient (SCG) algorithm was used to speed up the learning procedure in which the parameters of the Takagi–Sugeno–Kang fuzzy inference system (FIS) were tuned in the training set by shortening the training time per iteration. This method was referred to as neuro-fuzzy classifier (NFC) in this manuscript. In this FIS, “and,” “or,” “implication” from the antecedent to the consequent and “aggregation” of the consequents across the rules operators were product, probabilistic or “probor” (algebraic sum), minimum and maximum. Weighted defuzzification was also used. Note that Probor $(x, y) = x + y - x \times y$.

Feature selection

In many classification problems, a lot of candidate attributes are used for problem representation. Many of these are usually irrelevant or redundant.^[69] Thus, feature selection (FS) is used to detect relevant features usually leading to an increase in classifier accuracy.^[70] In this work, two supervised FS methodologies were used: (1) A statistical FS method, multiple logistic regression (MLR),^[71] and (2) a deterministic FS approach, sequential FS (SFS).^[72] MLR,

known as feature vector machine in machine learning, can be used to select statistically significant features.^[73] In our study, the intercept point was not used in the MLR. In SFS, the classifier starts with an empty set and added features until the accuracy was no longer improved by adding more features.

Validation

The performance of the classifier was assessed using the “hold-out” method, an approach to out-of-sample evaluation, in which the dataset was randomly split into two equal-size mutually exclusive sets (training and test sets). The classifier was trained on the training set and tested on the test set.^[74] The performance measures of the classification are listed in Table 2. The following performance measures were used: Sensitivity (Se), Specificity (Sp), Accuracy (Acc), and Precision (Pr) along with other indices reported in Table 2. Additionally, the McNemar’s (Gillick) statistical test was used^[75] to compare the performance of the NFC with different configurations (e.g., with or without FS) on the test set. Classification was performed using Matlab, Statistics Toolbox Release 2011a (The MathWorks, Inc., Natick, Massachusetts, USA) and the NFC toolbox.^[67,76] All statistical analyses and calculations were performed using the SPSS statistical package, version 18.0 (SPSS Inc., Chicago, IL, USA).

RESULTS

The number of subjects analyzed was 272 (68% male) out of total 303 in the Cleveland CAD dataset. The excluding criterion was the existence of missing value in any attribute. The characteristics of the raw dataset in the CAD and normal groups are shown in Table 3.

The performance of the NFC without FS, and with SFS/MLR on the preprocessed (discretized) Cleveland CAD dataset is shown in Table 4, in the hold-out validation framework. The feature set selected using SFS was (“thal,” “ST/HR slope,” “cp,” “ca,” “CHOL,” “trestbps,” “restecg,” “sex,” and “famhist”) from which the training accuracy increased from 72.1% to 88.2% during SFS procedure. Adding more features did not increase the accuracy. MLR, on the other hand, proposed the following five features

Table 2: The reported performance measures

Sensitivity (Se) = $\frac{TP}{TP + FN}$	Sensitivity (Se) = $\frac{TN}{TN + FP}$
Accuracy (Acc) = $\frac{TP+TN}{TP+TN+FP+FN}$	Precision(Pr) = $\frac{TP}{TP+FP}$
Recall (Rl)=Power=Se	F 1-score = $2 \times \frac{Pr \times Rl}{Pr + Rl}$
False alarm = $1 - Sp$ = False positive rate = α (type I error)	
$\beta = 1 - Se$ = False negative rate (type II error)	
TP = True positive; FN = False negative; FP = False positive; TN = True negative	

“age,” “exang,” “ca,” “thal,” and “ST/HR slope”). These features were statistically significant ($P < 0.05$) when running MLR in “Enter” mode excluding the intercept point in the model. Overall, MLR + NFC outperformed other classifiers in the training and test sets. It also required less input attributes for the decision-making procedure in comparison with those of other tested methods. McNemar’s test indicated that the performances of the SFS

+ NFC and MLR + NFC were higher in the training and test sets, respectively ($P < 0.05$). Since no tuning was performed in the test set and also because of required input attributes, MLR + NFC was chosen as the best classifier and used for further analysis.

The improvement of the classification error for MLR + NFC on the training set is shown in Figure 1. It depicted the fast reduction of root mean square error via SCG in the tuning procedure. The algorithm stopped at the iteration no. 100 since no significant improvement was made during the learning procedure.

Table 3: The attributes of the raw Cleveland dataset for normal and CAD groups, along with their categories (percentage) for nominal/ordinal variables and (minimum-maximum) mean ± SD for interval variables

Attribute*	Normal (n = 151)	CAD (n = 121)
Gender (%)	Male (54.8), female (45.2)	Male (84.0), female (16)
Cp (%)	1 (10.2), 2 (22.3), 3 (42.7), 4 (24.7)	1 (4.8), 2 (6.4), 3 (13.6), 4 (75.2)
FBS (%)	False (85.4), true (14.6)	False (84.8), true (15.2)
Famhist (%)	No (42.0), yes (58.0)	No (32.8), yes (67.2)
Restecg (%)	1 (56.7), 2 (0.6), 3 (42.7)	1 (39.2), 2 (0.8), 3 (60.0)
Exang (%)	No (85.4), yes (14.6)	No (44.8), yes (55.2)
Slope (%)	1 (64.3), 2 (30.6), 3 (5.1)	1 (27.2), 2 (64.8), 3 (8.0)
Thal (%)	1 (79.5), 2 (3.8), 3 (16.7)	1 (28.2), 2 (6.5), 3 (65.3)
Cigs	0-99 (15±19)	0-80 (17±20)
Age (year)	(29-76) 53±9	(35-77) 56±8
Trestbps (mmHg)	(94-180) 129±17	(100-200) 134±19
CHOL (mg/dL)	(126-564) 244±53	(149-409) 256±48
Years (year)	(0-50) 14±14	(0-54) 17±16
Thalach (bpm)	(96-202) 158±19	(71-195) 139±23
Thalrest (bpm)	(49-119) 77±14	(40-109) 73±13
Tpeakbps (mmHg)	(84-232) 170±23	(90-230) 165±25
Tpeakbpd (mmHg)	(26-120) 78±14	(50-120) 79±12
Trestbps (mmHg)	(50-110) 84±10	(68-110) 86±8
Oldpeak (µV)	(0-4) 0.6±0.8	(0-6) 1.6±1.3
Ca	(0-3) 1±1	(0-3) 1±1

*For the definition of attributes, their units and categories, the reader is referred to Table 1. CAD = Coronary artery disease; CHOL = Cholesterol; FBS = Fasting blood sugar; SD = Standard deviation

Table 4: The performance of the proposed NFC without FS, with SFS/MLR in the hold-out validation framework on the training and test sets

Scenario	Performance indices on the training set (%)				Performance indices on the test set (%)			
	Se	Sp	Pr	Acc	Se	Sp	Pr	Acc
NFC	88.5	85.3	83.1	86.8	81.7	79.0	75.4	80.2
SFS+NFC	88.5	88.0	85.7	88.2	81.7	77.6	74.2	79.4
MLR+NFC	75.4	85.3	80.7	80.9	81.7	92.1	89.1	87.5

SFS = Sequential feature selection; FS = Feature selection; MLR = Multiple logistic regression; NFC = Neuro-fuzzy classifier; Se = Sensitivity; Sp = Specificity; Pr = Precision; Acc = Accuracy. Scenarios: NFC = The proposed NFC without FS; SFS + NFC = Sequential FS + NFC; MLR + NFC = FS via multiple logistic regression + NFC

The MLR + NFC method that had been tuned on the training set was tested this time on the whole dataset for overall accuracy assessment. The resulting confusion matrix was shown in Table 5. The overall accuracy of the proposed classifier was 84.2%. Meanwhile, Cohen’s kappa coefficient, a statistical measure of inter-rater agreement, was 67.75% ($P < 0.05$) showing “substantial agreement”^[77] between the classifier’s outcome and that of coronary angiography.

The proposed fuzzy rule-based system (MLR + NFC) was shown in Figure 2. The system has two fuzzy rules. Each rule was related to an outcome diagnosis class. A diagnosis example was taken from the test set in the fuzzy rules to clarify the decision-making procedure. For a typical subject whose CAD was confirmed with angiography, the discretized attributes were as the following: Older adult age category, Angina was not induced by exercise, two major vessels colored by fluoroscopy, thallium-201 stress scintigraphy test showed reversible defect and ST/HR slope was high. These attributes were used as the input of the fuzzy system. The range of the output is within (0.9, 2.1) with the cut-off value of $1.5 = (0.9 + 2.1)/2$. If the output of the fuzzy system is higher or equal than 1.5, the subject has CAD. The output of the fuzzy system (based on the noninvasive tests) is 1.81, in agreement with that result of the gold standard (angiography). The first and second rules in Figure 2 have been extracted for normal and CAD subjects, respectively.

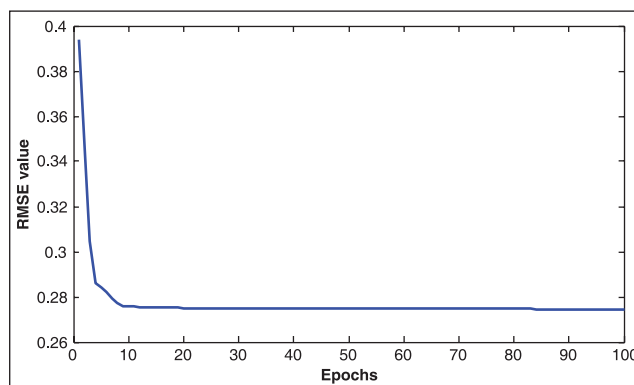


Figure 1: The improvement of the classification error for multiple logistic regression + neuro-fuzzy classifier on the training set versus epochs analyzed

DISCUSSION

Primary prevention, aiming at preventing heart and blood vessel disease in individuals who have not had a heart attack or symptoms of CAD and have no known clinical evidence of CAD, is highly recommended in comparison with secondary prevention for individuals with known CAD. This requires changing the lifestyle by quitting smoking, bringing down high blood pressure, controlling diabetes, maintaining a healthy body weight, doing regular exercises, following a heart-healthy diet to lower CHOL, LDL, and triglycerides, and to raise HDL, reducing stress and limiting alcohol consumption.^[8,78]

It has been suggested to integrate the prediction, prevention and intervention programs of CAD for more effective primary/secondary prevention.^[17,79] Invasive coronary angiography is the gold standard for establishing the presence, location, and severity of CAD.^[80,81] However, this technique is invasive, costly^[81] and associated with a small but definite risk of morbidity (1.5%) and mortality (0.15%).^[82,83] Therefore, a convenient, noninvasive alternative method for coronary angiography can provide significant clinical and economic benefits for the public health care system.^[84] A number of noninvasive CAD diagnosis methods have been proposed in the literature. Among which, fluoroscopy and stress thallium scintigraphy, are the two most popular. Unfortunately, the diagnostic accuracy of these methods in comparison with that of the gold standard, coronary angiography, ranges between 35% and 75%.^[85,86] Thus, the purpose of our work was to design a computer-aided noninvasive CAD diagnosis system using data mining methods to improve the diagnosis accuracy by combining the results of the noninvasive clinical tests and other attributes recorded from the subjects [Table 1].

A number of computer-aided CAD diagnosis systems have been proposed in the literature^[47,50,87-93] whose performance was tested on the Cleveland CAD dataset [Table 6]. Most of them use the black-box mathematical methodology that is not acceptable in medicine where the clinical interpretation of the decision-making procedure is critical. Our proposed fuzzy rule-based system, on the other hand, provided interpretable linguistic terms [Figure 2], which could be regarded as a fuzzy version of the decision tree classifier.^[94] For example, the comparison of rule 1 (normal) and rule 2 (CAD) in Figure 2 shows that the higher the ST/HR slope, and age, the higher the risk of CAD. Also, angina induced by exercise increases the risk of having CAD. Also, the last two categories of thallium-201 stress scintigraphy test (fixed and reversible defects) have higher CAD risk. The above rules are clinically acceptable. However, number of major vessels colored by fluoroscopy had a medium effect on the CAD diagnosis that could be due to the small sample size of the study population. It might be related to the fact that the sensitivity of fluoroscopy could be as low as 35% in some cases,^[86] and the system learned it from the training set. Meanwhile, the proposed system (MLR + NFC) only requires five input attributes for the diagnosis that is less than

Table 5: The overall confusion matrix of the MLR + NFC method*

MLR + NFC outcome	Patient with CAD confirmed with angiography	
	CAD positive	CAD negative
Test outcome positive	95 (TP)	17 (FP)
Test outcome negative	26 (FN)	134 (TN)

*The classifier was trained on the training set and tested on the whole dataset. "Positive" is related to "CAD diagnosis" while "negative" was used for "normal diagnosis". TP = True positive; FN = False negative; FP = False positive; MLR = Multiple logistic regression; NFC = Neuro-fuzzy classifier; CAD = Coronary artery disease; TN = True negative

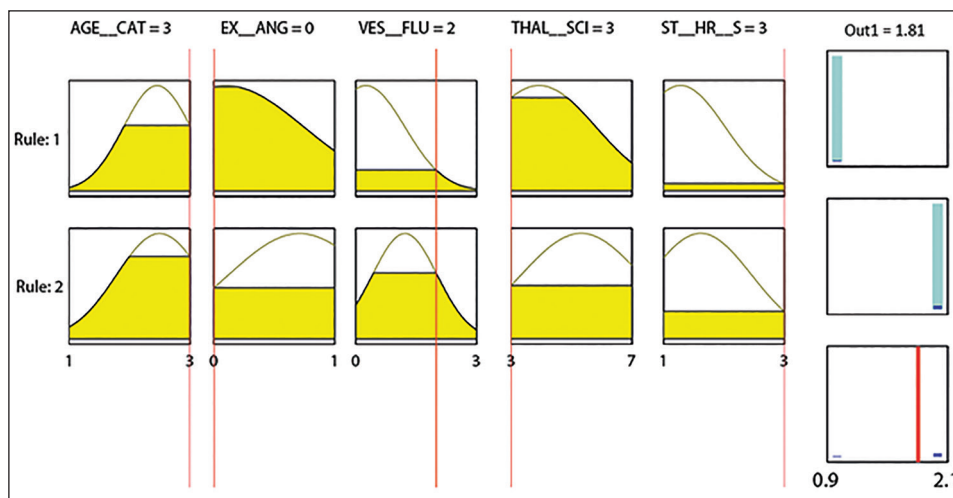


Figure 2: The extracted fuzzy rules from the training set. AGE_CAT: Age category, EX_ANG: Exercise-induced angina, VES_FLU: The number of vessels colored by fluoroscopy, THAL_SCI: Thallium-201 stress scintigraphy category, and ST_HR_S: Heart rate adjustment of exercise-induced ST segment depression category. Fuzzy rules 1 and 2 were related to normal and coronary artery diseases (CAD) classes, respectively. The overall output of the fuzzy system (1.81 in this example), is higher or equal to 1.5 indicating that the subject had CAD that was in agreement with what obtained from the gold standard (angiography). For the description of the input feature categories, refer to the section "preprocessing"

Table 6: Comparison of the proposed system outcome with similar research

Author	Method	Accuracy (%)
Detrano <i>et al.</i> , 1989	Probability theory (logistic regression)	77.00
Ster and Dobnikar, 1996	Fisher discriminant analysis	84.20
Ster and Dobnikar, 1996	LDA	84.50
Ster and Dobnikar, 1996	Naïve Bayes	82.50-83.40
Jankowski and Kadirkamanathan, 1997	IncNet	90.00
Cheung, 2001	BNNF	80.96
Cheung, 2001	C 4.5	81.11
Cheung, 2001	BNND	81.11
Cheung, 2001	Naïve Bayes	81.48
Polat <i>et al.</i> , 2007	Fuzzy-AIRS-Knn based system	87.00
Kahramanli and Allahverdi, 2008	Hybrid neural network system	86.80
Resul, 2009	Neural network ensembles	89.01
Senthil Kumar, 2011	ANFIS	91.18
Senthil Kumar, 2012	Fuzzy resolution mechanism	91.83
Muthukaruppan and Er, 2012	PSO based fuzzy expert system	93.27
Mahmoodabadi and Saniee Abadeh, 2014	Imperialist competitive algorithm based fuzzy expert system	94.92
Proposed classifier	MLR+NFC	84.00

ANFIS = Adaptive-network-based fuzzy inference system; MLR = Multiple logistic regression; NFC = Neuro-fuzzy classifier; PSO = Particle swarm optimization; AIRS = Artificial immune recognition system; LDA = Linear discriminant analysis

what has been proposed in the literature. Smaller number of required attributes facilitates the diagnosis procedure.

The overall accuracy of the proposed FIS (MLR + NFC) was 84% [Table 5]. Since the number of normal subjects in the dataset was higher than CAD [Table 3], the F_1 -score was also calculated (=82%) since it is an unbiased accuracy measure in unbalanced datasets. The type I error (α) and the power of the proposed diagnosis test were 0.1 and 76%. Reducing α down to 0.05 and increasing power to 80%, improves the clinical reliability of the proposed system that is the focus of our future activity. The following strategies could be taken in this regard:

1. The selected mixed-type features in MLR could be weighted based on the clinical importance and (or) tuned using particle swarm optimization^[95] embedded with Generalized Minkowski Metrics.^[96] Tuning the feature weights, might improve the performance of the designed classifier;
2. Although the sample size of the Cleveland dataset is not low, some important features such as body mass index (BMI) are missing.

We are thus going to design an automated CAD risk assessment program, based on the findings of this study, in collaboration with Isfahan Healthy Heart Program^[97] whose dataset is quite rich.

In addition to the hold-out validation method, we also used a 10-fold cross-validation in which the original sample was randomly partitioned into 10 equal-size subsamples. Of the 10 subsamples, a single subsample was retained as the validation data for testing the model, and the remaining 9 subsamples were used as training data. The cross-validation process was then repeated 10 times (the folds), with each of the 10 subsamples used exactly once as the validation data. The results from the folds were averaged to estimate a single estimation.^[98] The overall cross-validation accuracy of the proposed classifier was 83%. Guarding against testing hypotheses suggested by the data (type III errors^[99]) was done by cross-validation. Comparison with other diagnosis methods designed on the Cleveland CAD dataset, our proposed method ranked in the top first quartile [Table 6].

Although the BMI was not used in our model, it is well-known that BMI is positively correlated with CHOL.^[100-102] In our dataset, age was positively correlated with hypertension, CHOL, and FBS. This might explain that age in the final model could capture information about CHOL, hypertension, FBS, and BMI as well. The number of cigarettes per day and number of years as a smoker were positively correlated with thallium-201 stress scintigraphy categories. Thus, this attribute could capture the smoking information. Gender was associated with the exercise-induced angina status. Thus, information about the gender could be captured by exercise-induced angina. This is, in fact, the property of MLR that takes into account the interaction between attributes and reports the compact attribute set. Accordingly, most of the input risk factors could be taken into accounts directly/indirectly. Among the features selected, the following ones have been selected in another manuscript:^[103] "ST/HR slope," "cp," "age," "trestbps."

CONCLUSION

We proposed an interpretable fuzzy rule-based system that could noninvasively predict the CAD based on "age," "exercise-induced angina status," number of major vessels colored by fluoroscopy, thallium-201 stress scintigraphy result and ST/HR slope. The proposed computer-aided system was promising in CAD diagnosis and could be implemented as a web-based diagnostic decision support system. However, its performance could be improved by introducing weights to the input attributes taking into account the clinical relevance/priority of the features.

ACKNOWLEDGMENT

This work was supported by the University of Isfahan.

AUTHOR'S CONTRIBUTION

HRM contributed in the conception and design of the work, drafting and revising the draft, approval of the final version of the manuscript, and agreed for all aspects of the work. SG contributed in the conception of the work, revising the draft, approval of the final version of the manuscript, and agreed for all aspects of the work.

REFERENCES

- Libby P, Theroux P. Pathophysiology of coronary artery disease. *Circulation* 2005;111:3481-8.
- Simon S. *The Heart: Our Circulatory System*. New York: Morrow Junior Books; 1996.
- McGoon MD. *Mayo Clinic Heart Book*. New York: W. Morrow; 1993.
- Khatibi V, Montazer GA. A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment. *Expert Syst Appl* 2010;37:8536-42.
- Hatmi ZN, Tahvildari S, Gafarzadeh Motlag A, Sabouri Kashani A. Prevalence of coronary artery disease risk factors in Iran: A population based survey. *BMC Cardiovasc Disord* 2007;7:32.
- Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Baha MJ, *et al.* Executive summary: Heart disease and stroke statistics-2014 update: A report from the American Heart Association. *Circulation* 2014;129:399-410.
- Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, *et al.* Executive summary: Heart disease and stroke statistics -2013 update: A report from the American Heart Association. *Circulation* 2013;127:143-52.
- Barnes GW, Holmes DR. *Coronary Artery Disease: New Approaches without Traditional Revascularization*. New York, London: Springer; 2008. p. 187.
- Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson TB, Flegal K, *et al.* Heart disease and stroke statistics-2009 update: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2009; 119:480-6.
- Mendis S, Puska P, Norrving B. World Health Organization, World Heart Federation, World Stroke Organization, Global atlas on cardiovascular disease prevention and control. World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization: Geneva, 2011; p. vi, p. 155.
- Mensah GA. *The Atlas of Heart Disease and Stroke*. Geneva: World Health Organization; 2004.
- Ebrahimi M, Kazemi-Bajestani SM, Ghayour-Mobarhan M, Ferns GA. Coronary artery disease and its risk factors status in Iran: A review. *Iran Red Crescent Med J* 2011;13:610-23.
- Hadaegh F, Harati H, Ghanbarian A, Azizi F. Prevalence of coronary heart disease among Tehran adults: Tehran Lipid and Glucose Study. *East Mediterr Health J* 2009;15:157-66.
- Sarraf-Zadegan N, Sayed-Tabatabaei FA, Bashardoost N, Maleki A, Totonchi M, Habibi HR, *et al.* The prevalence of coronary artery disease in an urban population in Isfahan, Iran. *Acta Cardiol* 1999;54:257-63.
- Foody JM. *Preventive Cardiology: Strategies for the Prevention and Treatment of Coronary Artery Disease*, Contemporary Cardiology. Totowa, NJ: Humana Press; 2001. p. x, 366.
- McMahan CA, Gidding SS, Fayad ZA, Zieske AW, Malcom GT, Tracy RE, *et al.* Risk scores predict atherosclerotic lesions in young people. *Arch Intern Med* 2005;165:883-90.
- Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837-47.
- Okuyama H. *Prevention of Coronary Heart Disease: From the Cholesterol Hypothesis to [Omega]6/ [Omega]3 Balance*, World Review of Nutrition and Dietetics. Basel, New York: Karger; 2007. p. x, 168.
- Mittal S. *Coronary Heart Disease in Clinical Practice*. London: Springer; 2005.
- Li D, Qu C, Dong P. The ICAM-1 K469E polymorphism is associated with the risk of coronary artery disease: A meta-analysis. *Coron Artery Dis* 2014;25:665-70.
- Barth J, Schumacher M, Herrmann-Lingen C. Depression as a risk factor for mortality in patients with coronary heart disease: A meta-analysis. *Psychosom Med* 2004;66:802-13.
- Kivimäki M, Nyberg ST, Fransson EI, Heikkilä K, Alfredsson L, Casini A, *et al.* Associations of job strain and lifestyle risk factors with risk of coronary artery disease: A meta-analysis of individual participant data. *CMAJ* 2013;185:763-9.
- Roest AM, Martens EJ, de Jonge P, Denollet J. Anxiety and risk of incident coronary heart disease: A meta-analysis. *J Am Coll Cardiol* 2010;56:38-46.
- Dauchet L, Amouyel P, Hercberg S, Dallongeville J. Fruit and vegetable consumption and risk of coronary heart disease: A meta-analysis of cohort studies. *J Nutr* 2006;136:2588-93.
- Cahill LE, Pan A, Chiuve SE, Sun Q, Willett WC, Hu FB, *et al.* Fried-food consumption and risk of type 2 diabetes and coronary artery disease: A prospective study in 2 cohorts of US women and men. *Am J Clin Nutr* 2014;100:667-675.
- Castellino RA. Computer aided detection (CAD): An overview. *Cancer Imaging* 2005;5:17-9.
- Rangayyan RM. *Biomedical Signal Analysis: A Case-Study Approach*. IEEE Press Series in Biomedical Engineering. Piscataway, NJ, New York, NY: IEEE Press, Wiley-Interscience; 2002. p. xxxv, 516.
- Rogers W, Ryack B, Moeller G. Computer-aided medical diagnosis: Literature review. *Int J Biomed Comput* 1979;10:267-89.
- Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: A systematic review. *JAMA* 1998;280:1339-46.
- Salas-Gonzalez D, Górriz JM, Ramírez J, López M, Alvarez I, Segovia F, *et al.* Computer-aided diagnosis of Alzheimer's disease using support vector machines and classification trees. *Phys Med Biol* 2010;55:2807-17.
- Alpaydin E. *Adaptive computation and machine learning*. Introduction to Machine Learning. 2nd ed. Cambridge, Mass: MIT Press; 2010. p. xl, 537.
- Duda RO, Hart PE, Stork DG. *Pattern Classification*. New York: John Wiley & Sons; 2012.
- Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang JF, *et al.* Data mining in healthcare and biomedicine: A survey of the literature. *J Med Syst* 2012;36:2431-48.
- Esfandiari N, Babavalian MR, Moghadam AM, Tabar VK. Review: Knowledge discovery in medicine: Current issue and future trend. *Expert Syst Appl* 2014;41:4434-63.
- Mykowiecka A, Kupś A, Marciniak M. Rule-based medical content extraction and classification. In: Kłopotek M, Wierchoń S, Trojanowski K, editors. *Intelligent Information Processing and Web Mining*. New York: Springer Berlin Heidelberg; 2005. p. 237-45.

36. Ruspini EH, Bonissone PP, Pedrycz W. Handbook of fuzzy computation. Institute of Physics Pub: Philadelphia, 1998; p. G13.
37. Innocent PR, John RI. Computer aided fuzzy medical diagnosis. *Inf Sci* 2004;162:81-104.
38. Adlassnig KP. Fuzzy set theory in medical diagnosis. *Syst Man Cybern IEEE Trans* 1986;16:260-5.
39. Seising R. From vagueness in medical thought to the foundations of fuzzy reasoning in medical diagnosis. *Artif Intell Med* 2006;38:237-56.
40. Teodorescu HN, Kandel A, Jain LC. Fuzzy and Neuro-Fuzzy Systems in Medicine. Vol. 2. Boca Raton: CRC Press; 1998.
41. Gadaras I, Mikhailov L. An interpretable fuzzy rule-based classification methodology for medical diagnosis. *Artif Intell Med* 2009;47:25-41.
42. Nauck D, Kruse R. Obtaining interpretable fuzzy classification rules from medical data. *Artif Intell Med* 1999;16:149-69.
43. Sanz JA, Galar M, Jurio A, Brugos A, Pagola M, Bustince H. Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system. *Appl Soft Comput* 2014;20:103-11.
44. Olaru C, Wehenkel L. A complete fuzzy decision tree technique. *Fuzzy Sets Syst* 2003;138:221-54.
45. Wang T, Li Z, Yan Y, Chen H. A survey of fuzzy decision tree classifier methodology. In: Cao BY, editor. *Fuzzy Information and Engineering*. New York: Springer Berlin Heidelberg; 2007. p. 959-68.
46. Lichman M. UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. University of California, School of Information and Computer Science: Irvine, CA, 2013. Access Date: Feb, 2014.
47. Detrano R, Janosi A, Steinbrunn W, Pfisterer M, Schmid JJ, Sandhu S, *et al.* International application of a new probability algorithm for the diagnosis of coronary artery disease. *Am J Cardiol* 1989;64:304-10.
48. Aha D, Kibler D. Instance-based prediction of heart-disease presence with the Cleveland database. Technical Report, University of California, Irvine, Department of Information and Computer Science, Report Number ICS-TR-88-07. 1988.
49. Gennari JH, Langley P, Fisher D. Models of incremental concept formation. *Artif Intell* 1989;40:11-61.
50. Muthukaruppan S, Er MJ. A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. *Expert Syst Appl* 2012;39:11657-65.
51. Detrano R, Yiannikas J, Salcedo EE, Rincon G, Go RT, Williams G, *et al.* Bayesian probability analysis: A prospective demonstration of its clinical utility in diagnosing coronary disease. *Circulation* 1984;69:541-7.
52. Setiawan NA, Venkatachalam P, Hani AF. Diagnosis of Coronary Artery Disease Using Artificial Intelligence Based Decision Support System. In Proceedings of the International Conference on Man-Machine Systems (ICoMMS), Batu Ferringhi, Penang; 2009.
53. Lustgarten JL, Gopalakrishnan V, Grover H, Visweswaran S. Improving classification performance with discretization on biomedical datasets. *AMIA ... Annual Symposium proceedings / AMIA Symposium*. AMIA Symposium 2008, p. 445-9.
54. Liu H, Setiono R. Feature selection via discretization. *IEEE Trans Knowl Data Eng* 1997;9:642-5.
55. Kotsiantis S, Kanellopoulos D. Discretization techniques: A recent survey. *GESTS Int Trans Comput Sci Eng* 2006;32:47-58. Available from: <http://www.gests.org/>
56. Cangelosi D, Muselli M, Parodi S, Blengio F, Becherini P, Versteeg R, *et al.* Use of Attribute Driven Incremental Discretization and Logic Learning Machine to build a prognostic classifier for neuroblastoma patients. *BMC Bioinformatics* 2014;15 Suppl 5:S4.
57. Wong DF, Chao LS, Zeng XD. A supportive attribute-assisted discretization model for medical classification. *Biomed Mater Eng* 2014;24:289-95.
58. Petry NM. A comparison of young, middle-aged, and older adult treatment-seeking pathological gamblers. *Gerontologist* 2002;42:92-9.
59. Hsia J, Margolis KL, Eaton CB, Wenger NK, Allison M, Wu L, *et al.* Prehypertension and cardiovascular disease risk in the Women's Health Initiative. *Circulation* 2007;115:855-60.
60. Yamaguchi N, Mochizuki-Kobayashi Y, Utsunomiya O. Quantitative relationship between cumulative cigarette consumption and lung cancer mortality in Japan. *Int J Epidemiol* 2000;29:963-8.
61. Fox K, Borer JS, Camm AJ, Danchin N, Ferrari R, Lopez Sendon JL, *et al.* Resting heart rate in cardiovascular disease. *J Am Coll Cardiol* 2007;50:823-30.
62. Kligfield P, Ameisen O, Okin PM, Borer JS. Evaluation of the exercise electrocardiogram by the ST segment/heart rate slope. *Bull N Y Acad Med* 1987;63:480-92.
63. Okin PM, Grandits G, Rautaharju PM, Prineas RJ, Cohen JD, Crow RS, *et al.* Prognostic value of heart rate adjustment of exercise-induced ST segment depression in the multiple risk factor intervention trial. *J Am Coll Cardiol* 1996;27:1437-43.
64. Mitra S, Hayashi Y. Neuro-fuzzy rule generation: Survey in soft computing framework. *IEEE Trans Neural Netw* 2000;11:748-68.
65. Jang JS, Sun CT, Mizutani E. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, MATLAB Curriculum Series. Upper Saddle River, NJ: Prentice Hall; 1997. p. xxvi, 614.
66. Jang JS. Anfis: Adaptive-Network-Based Fuzzy Inference System. *Syst Man Cybern IEEE Trans* 1993;23:665-85.
67. Cetişli B, Barkana A. Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy classifier training. *Soft Comput* 2010;14:365-78.
68. Sun CT, Jang JS. A Neuro-Fuzzy Classifier and Its Applications. In *Fuzzy Systems, 1993, Second IEEE International Conference on IEEE*; 1993. p. 94-8.
69. Dash M, Liu H. Feature selection for classification. *Intell Data Anal* 1997;1:131-56.
70. Marateb HR, Mansourian M, Faghihimani E, Amini M, Farina D. A hybrid intelligent system for diagnosing microalbuminuria in type 2 diabetes patients without having to measure urinary albumin. *Comput Biol Med* 2014;45:34-42.
71. Hellevik O. Linear versus logistic regression when the dependent variable is a dichotomy. *Qual Quant* 2009;43:59-74.
72. Jain A, Zongker D. Feature selection: Evaluation, application, and small sample performance. *Pattern Anal Mach Intell IEEE Trans* 1997;19:153-8.
73. Li F, Yang Y, Xing EP. From lasso regression to feature vector machine. In: Weiss Y, editor. *Advances in Neural Information Processing Systems*. Proceedings of a meeting held 5-8 December 2005, Vancouver, British Columbia, Canada; Neural Information Processing Systems (NIPS). 2005. p. 779-86.
74. Sammut C, Webb GI. *Encyclopedia of Machine Learning*. New York: Springer; 2011.
75. Dietterich TG. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput* 1998;10:1895-923.
76. Cetişli B. Neuro-Fuzzy Classifier Toolbox. 2010. Available from: <http://www.mathworks.com/matlabcentral/fileexchange/29043-neuro-fuzzy-classifier>. [Last accessed on 2014 Feb].
77. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
78. World Health Organization, UNAIDS. *Prevention of Cardiovascular Disease*. Geneva: World Health Organization; 2007.
79. Grover SA, Coupal L, Hu XP. Identifying adults at increased risk of coronary disease. How well do the current cholesterol guidelines work? *JAMA* 1995;274:801-6.

80. Levin DC. Invasive evaluation (coronary arteriography) of the coronary artery disease patient: Clinical, economic and social issues. *Circulation* 1982;66:III71-9.
81. Ryan TJ. The coronary angiogram and its seminal contributions to cardiovascular medicine over five decades. *Circulation* 2002;106:752-6.
82. Adams DF, Fraser DB, Abrams HL. The complications of coronary arteriography. *Circulation* 1973;48:609-18.
83. Kennedy JW. Complications associated with cardiac catheterization and angiography. *Cathet Cardiovasc Diagn* 1982;8:5-11.
84. Kohsaka S, Makaryus AN. Coronary angiography using noninvasive imaging techniques of cardiac CT and MRI. *Curr Cardiol Rev* 2008;4:323-30.
85. Detrano R, Simpfendorfer C, Day K, Salcedo EE, Rincon G, Kramer JR, *et al.* Comparison of stress digital ventriculography, stress thallium scintigraphy, and digital fluoroscopy in the diagnosis of coronary artery disease in subjects without prior myocardial infarction. *Am J Cardiol* 1985;56:434-40.
86. Vliegenthart R. Detection and quantification of coronary calcification. In: Oudkerk M, editor. *Coronary Radiology*. New York: Springer Berlin Heidelberg; 2004. p. 175-84.
87. Šter B, Dobnikar A. Neural Network in Medical Diagnosis: Comparison with Other Methods. In *International Conference on Engineering Applications of Neural Networks*; 1996. p. 427-30.
88. Cheung N. *Machine Learning Techniques for Medical Analysis*. School of Information Technology and Electrical Engineering, B. Sc. Thesis, University of Queensland; 2001.
89. Kahramanli H, Allahverdi N. Design of a hybrid system for the diabetes and heart diseases. *Expert Syst Appl* 2008;35:82-9.
90. Polat K, Şahan S, Güneş S. Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and K -Nn (nearest neighbour) based weighting preprocessing. *Expert Syst Appl* 2007;32:625-31.
91. Das R, Turkoglu I, Sengur A. Effective diagnosis of heart disease through neural networks ensembles. *Expert Syst Appl* 2009;36:7675-80.
92. Jankowski N, Kadirkamanathan V. Statistical control of RBF-like networks for classification. In: *Artificial Neural Networks-ICANN'97*. Lausanne: Springer; 1997. p. 385-90.
93. Senthil Kumar AV. Diagnosis of Heart Disease using Fuzzy Resolution Mechanism. *Journal of Artificial Intelligence*, 2012;5: 47-55. DOI: 10.3923/jai.2012.47.55.
94. Yuan Y, Shaw MJ. Induction of fuzzy decision trees. *Fuzzy Sets Syst* 1995;69:125-39.
95. Marateb HR, McGill KC. Resolving superimposed MUAPs using particle swarm optimization. *IEEE Trans Biomed Eng* 2009;56:916-9.
96. Ichino M, Yaguchi H. Generalized minkowski metrics for mixed feature-type data analysis. *Syst Man Cybern IEEE Trans* 1994;24:698-708.
97. Sarraf-Zadegan N, Sadri G, Malek Afzali H, Baghaei M, Mohammadi Fard N, Shahrokhi S, *et al.* Isfahan Healthy Heart Programme: A comprehensive integrated community-based programme for cardiovascular disease prevention and control. Design, methods and initial experience. *Acta Cardiol* 2003;58:309-20.
98. Geisser S. *Predictive Inference*. Vol. 55. New York: CRC Press; 1993.
99. Mosteller F. AK-sample slippage test for an extreme population. In: *Selected Papers of Frederick Mosteller*. New York: Springer; 2006. p. 101-9.
100. Faheem M, Qureshi S, Ali J, Hameed, Zahoor, Abbas F, *et al.* Does BMI affect cholesterol, sugar, and blood pressure in general population? *J Ayub Med Coll Abbottabad* 2010;22:74-7.
101. Gostynski M, Gutzwiller F, Kuulasmaa K, Döring A, Ferrario M, Grafnetter D, *et al.* Analysis of the relationship between total cholesterol, age, body mass index among males and females in the WHO MONICA Project. *Int J Obes Relat Metab Disord* 2004;28:1082-90.
102. Lamon-Fava S, Wilson PW, Schaefer EJ. Impact of body mass index on coronary heart disease risk factors in men and women. The Framingham Offspring Study. *Arterioscler Thromb Vasc Biol* 1996;16:1509-15.
103. Alizadehsani R, Habibi J, Hosseini MJ, Mashayekhi H, Boghrati R, Ghandeharioun A, *et al.* A data mining approach for diagnosis of coronary artery disease. *Comput Methods Programs Biomed* 2013;111:52-61.

Source of Support: Nil, **Conflict of Interest:** None declared.