

Dietary protein is the strong predictor of coronary artery disease; a data mining approach

Sara Saffar Soflaei

Mashhad University of Medical Sciences

Elham Shamsara

Mashhad University of Medical Sciences

Toktam Sahranavard

Mashhad University of Medical Sciences

Habibollah Esmaily

Mashhad University of Medical Sciences

Mohsen Moohebat

Mashhad University of Medical Sciences

Niloofer Shabani

Mashhad University of Medical Sciences

Zahra Asadi

Mashhad University of Medical Sciences

Mohammad Tajfard

Mashhad University of Medical Sciences

Gordon A Ferns

Brighton and Sussex Medical School

Majid Ghayour Mobarhan (✉ ghayourm@mums.ac.ir)

Mashhad University of Medical Sciences <https://orcid.org/0000-0002-1081-6754>

Research

Keywords: coronary artery disease, FFQ, Quest model, dietary intake, dietary protein, data mining

Posted Date: August 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-57487/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at Clinical Nutrition ESPEN on June 1st, 2021.
See the published version at <https://doi.org/10.1016/j.clnesp.2021.03.008>.

Abstract

Backgrounds and aims: Coronary artery disease (CAD) is the major cause of mortality and morbidity globally. Diet is known to contribute to CAD risk, and the dietary intake of specific macro- or micro-nutrients might be potential predictors of CAD risk. Machine learning methods may be helpful in the analysis of the contribution of several parameters in dietary including macro- and micro-nutrients to CAD risk. Here we aimed to determine the most important dietary factors for predicting CAD.

Methods: Total 273 cases with more than 50% obstruction in at least one coronary artery and 443 healthy controls who completed a food frequency questionnaire (FFQ) were entered into the study. All dietary intakes were adjusted for energy intake. QUEST method was applied to determine the diagnosis pattern of CAD.

Results: Total 34 dietary variables obtained from FFQ were entered the study that 23 of these variables were significantly associated with CAD according to t-test. Out of 23 dietary input variables adjusted protein, manganese, biotin, zinc and cholesterol remained in the model. According to our tree, only protein intake could identify the patients with coronary artery stenosis according to angiography from healthy participant up to 80%. Manganese dietary intake was the second important variable after protein. The accuracy of the tree was 84.36% for training dataset and 82.94% for testing dataset.

Conclusion: Among different macro- and micro-nutrients in the dietary, a combination of protein, manganese, biotin, zinc and cholesterol could predict the presence of CAD.

Introduction

Coronary artery disease (CAD) is the major cause of morbidity and mortality worldwide (1). CAD is also very prevalent in Iran compared to other countries. Therefore, finding a national program for reducing the risk factors of CAD based on lifestyle is fundamental (2). The most prevalent CAD risk factors are smoking, male gender, age, ethnicity, family history of the disease, high blood pressure, high blood cholesterol, diabetes, poor diet, Lack of exercise, obesity, stress and blood vessel inflammation. These factors affect each patient variously (3).

The gold standard of CAD diagnosis remains invasive coronary angiography, but this procedure is associated with a risk of serious complication (4). Finding an appropriate, safe and non-invasive method to diagnosis is the aim of current diagnostic approaches (4). Evidence indicated significant association of a limited number of dietary factors and dietary patterns with CAD (5). Previous association studies reported mineral dietary intake such as sodium, potassium, magnesium and zinc as associated risk factors of CAD (6–11). While, many previous studies have shown that vitamin C, vitamin E and selenium interventions do not reduce the risk of CAD (12, 13). Furthermore, a study in China indicated that low fat and high fiber intake decrease CAD mortality (14). In regard to protein intake and risk of CAD, in a follow-up study by health professionals a significant relationship was found between total protein and increased

risk of CAD (15), However, in a review article, Pedersen et al. concluded that there was no significant relationship between protein intake and strokes and coronary heart disease (16).

Hence the use of dietary and intake patterns and their application to novel algorithms to predict CAD remains a substantial approach to risk stratification (17).

Machine learning is rapidly used to predict healthcare issues, such as cost, utilization, and status. In machine learning, the purpose is to train the algorithm to learn how maps inputs features to an output. Generally, any machine learning method applies the following steps; data preparation, algorithm selection, training, regularization, and evaluation (18). Different methods of machine learning models for coronary artery disease were previously built and analyzed (19–21). Nevertheless, the circumstances may vary based on different situations, lifestyles, and accessible data and features. Thus, we believe that with constructing and validating prediction models, it becomes plausible to classify patients who have a high risk of disease from those who are at low risk. Consequently, a diagnostic model for predicting CAD is necessary.

In the present paper, among different methods of machine learning (artificial neural network, deep learning, etc.) we employed a well-known technique called decision tree (DT). A DT model is a graphical model that its structure is like a tree. One of the advantage of DT is that the produced model is a more interpretable model.

DT is a predictive technique which turns the fact about a disease into some conclusions about the disease purpose value. DT is one of the most significant algorithms applied in machine learning. C5.0, C&R Tree, CHAID, and QUEST (Quick, Unbiased, Efficient Statistical Tree) are some applying DT algorithms in machine learning modeling.

QUEST is a binary-split decision tree method of machine learning. In Quest, the association between the input features and the target is calculated by ANOVA F-test (ordinal features) or Pearson's chi-square (nominal features). The features that make the greatest agreement with the target is chosen to divide the node (22). The computation speed in this method is greater than those in other algorithms; the benefit of this algorithm is that it can avoid the bias that exists in other classification methods (23).

In this current study, QUEST is applied for models construction to recognize the importance of factors related to incidence of CAD, and detecting dietary intake as a major CAD risk factor.

Materials And Methods

Subjects

The data was extracted from our previous case-control study, between September 2011 and May 2013 (17). Out of 1187 patients whom underwent coronary angiography, 273 cases whom had more than 50%

obstruction in at least one coronary artery and also their food frequency questionnaire (FFQ) was available, entered the study.

Healthy controls were selected from the same study. The healthy subjects had no signs or symptoms of CAD. Furthermore, they did not have any of the traditional risk factors of CAD. Total of 443 healthy controls who had FFQ questionnaire were chosen.

FFQ

The dietary intake data of the current study population were collected by a semi-quantitative food frequency questionnaire (FFQ) which was validated among an Iranian population (19). This FFQ is a 65-food item one and each food item was consisted of frequency intake (per day, per week, per month, seldom, and never) and portion size. After completing the FFQ by experienced nutritionists, dietary intake data was analyzed by diet plan 7 software. Consequently, dietary intake of micronutrients and macronutrients was obtained for all subjects.

Data adjustment

We performed an energy adjustment method for each input attributes. We applied the energy adjustment method based on the residual. In this method, the energy-adjusted intake measure is the residual from a regression model in which total energy intake is the independent variable and absolute nutrient intake is the dependent variable (2, 24, 25).

All the variables which were significant between participants with positive angiography and healthy participants were considered as input variables. The input variables were adjusted protein, carbohydrate, sugar, fiber, total fat, cholesterol, mono unsaturated fat, sodium, potassium, phosphorus, calcium, magnesium, iodine, manganese, zinc, selenium, carotene, folate, vitamin C, thiamin, retinol, niacin, biotin shown in Table 1. The model evaluated in this study had 10 input variables and one target variable. The target variable consisted of 2 classes as healthy and positive angiography.

Table 1
Comparison of dietary intakes between healthy and positive angiography.

Dietary intake ^a		Healthy		Angiogram +		p value (t test)
		<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	
Macronutrients	Protein (g/day)	58.99	22.18	100.43	37.68	< 0.001
	Carbohydrate (g/day)	241.99	41.64	219.84	43.15	< 0.001
	Starch (g/day)	151.44	42.29	153.40	36.53	0.795
	Sugar (g/day)	89.15	44.62	63.89	44.62	< 0.001
	Fiber (g/day)	15.46	7.30	11.22	7.27	< 0.001
	Total fat (g/day)	68.82	15.15	60.80	14.47	0.023
	Cholesterol (mg/day)	207.22	153.72	220.41	118.75	0.001
	Saturated fat (g/day)	17.25	6.28	17.39	5.81	0.950
	MUFA (g/day)	17.14	4.96	15.10	4.83	< 0.001
	PUFA (g/day)	21.12	8.75	21.15	8.82	0.999
	Sodium (mg/day)	1845.99	1706.37	1426.27	687.34	0.025
	Potassium (mg/day)	2432.98	1006.40	3122.37	1506.17	< 0.001
	Phosphorus (mg/day)	1240.50	288.26	1323.90	315.89	0.001
	Calcium (mg/day)	809.11	281.07	871.37	378.40	0.036
Micronutrients	Magnesium (mg/day)	230.34	76.58	233.88	53.37	0.034
	Iodine (µg/day)	97.17	66.33	104.92	57.88	0.009
	Iron (mg/day)	10.46	4.31	11.61	9.04	0.307
	Manganese (mg/day)	3.22	1.45	2.18	0.93	< 0.001
	Copper (µg/day)	1.22	1.11	1.22	1.09	0.953
	Zinc (mg/day)	8.23	2.54	9.46	3.11	< 0.001
	Selenium (µg/day)	34.98	22.93	49.63	29.75	< 0.001
	Carotene (µg/day)	2057.68	2096.05	1975.62	4201.60	< 0.001
	Folate (µg/day)	224.81	151.66	186.74	91.17	< 0.001

Abbreviations: MUFA: mono unsaturated fatty acid; PUFA: poly unsaturated fatty acid.

^a All nutrients were adjusted for total energy intake.

Vitamin E (mg/day)	15.03	8.23	15.72	7.42	0.475
Vitamin D (µg/day)	1.43	1.84	2.49	4.17	0.282
Vitamin C (mg/day)	83.95	68.26	68.94	92.52	< 0.001
Thiamin (mg/day)	1.31	0.49	1.35	0.40	0.029
Retinol (mg/day)	382.68	2187.35	369.84	1883.14	< 0.001
Riboflavin (mg/day)	1.33	0.87	1.38	1.04	0.476
Vitamin B12 (µg/day)	3.63	8.66	3.89	10.67	0.710
Vitamin B6 (mg/day)	1.47	0.97	1.53	0.60	0.283
Niacin (mg/day)	14.73	5.57	24.74	13.21	< 0.001
Pantothenic acid (mg/day)	4.83	3.30	5.19	2.21	0.080
Biotin (µg/day)	31.77	42.99	18.26	10.99	< 0.001

Abbreviations: MUFA: mono unsaturated fatty acid; PUFA: poly unsaturated fatty acid.

^a All nutrients were adjusted for total energy intake.

Model

In this model, the QUEST method has been investigated to analyze the data and build a diagnosis pattern of patients with coronary artery disease. To perform the investigation, total number of 716 participants were considered. As a common rule in decision tree, data were divided into training and testing groups, 70% of total participants (505 subjects) were randomly selected to make training group for constructing the decision tree. The remaining 30% (211 subjects) were considered as testing group to evaluate the performance of decision tree.

A confusion matrix was used to evaluate the performance of the decision-tree for classification of participants. The accuracy, sensitivity, specificity and the receiver operating characteristics (ROC) curve were measured for comparison.

Results

Micro and macronutrients obtained according to FFQ questionnaire for total number of 716 subjects in two groups of angiogram positive and healthy subjects were indicated in Table 1. Out of 23 input variables adjusted protein, manganese, biotin, zinc and cholesterol remained in the model. The final decision tree with 12 leaves and 4 layers was shown in Fig. 1. The if-then rules is shown in Table 2.

Table 2
The if-then rules extracted from the tree

Rule 1: If protein \leq 80.076 g/day, Manganese mg/day \leq 2.305, biotin \leq 17.985 μ g/day and cholesterol \leq 168.181 mg/day, Then class: A+ (22/23 or 95.65%)
Rule 2: If protein \leq 80.076 g/day, Manganese mg/day \leq 2.305, biotin \leq 17.985 μ g/day and cholesterol $>$ 168.181 mg/day, Then class: H (9/11 or 81.82%)
Rule 3: If protein \leq 80.076, Manganese \leq 2.305 mg/day and biotin $>$ 17.985 μ g/day, Then class: H (34/41 or 82.93%)
Rule 4: If protein \leq 80.076 g/day and Manganese $>$ 2.305 mg/day, Then class: H (134/146 or 91.70%)
Rule 5: If protein $>$ 80.076 g/day, Manganese \leq 2.870 mg/day and biotin \leq 36.376 μ g/day, Then class: A+ (97/103 or 94.18%)
Rule 6: If protein $>$ 80.076 g/day, Manganese \leq 2.870 mg/day and biotin $>$ 3.376 μ g/day, Then class: H (5/6 or 83.33%)
Rule 7: If protein $>$ 80.076 g/day and Manganese $>$ 2.870 mg/day, Then class: H (22/36 or 61.11%)
H, healthy; A+, angiography positive.

For evaluation of the decision tree, confusion matrix were used which was indicated in Table 3 for training and testing datasets. The accuracy of the tree was 84.36% for training dataset and 82.94% for testing dataset. Other performance variables of the tree including sensitivity, specificity and AUC was shown in Table 4.

Table 3
Confusion matrix of training and testing datasets.

Training		Testing			
Actual outcome	Predicted outcome		Actual outcome	Predicted outcome	
	<i>H</i>	<i>A +</i>		<i>H</i>	<i>A +</i>
<i>H</i>	284	18	<i>H</i>	133	8
<i>A +</i>	61	142	<i>A +</i>	28	42
H, healthy; A+, angiography positive.					

Table 4

The performance of the Quest decision tree to identify associated risk factors of CAD (for test and train datasets)

Variables	Quest model	
Sensitivity	<i>Training</i>	82.3%
	<i>Testing</i>	82.61%
Specificity	<i>Training</i>	88.75%
	<i>Testing</i>	84.00%
Accuracy	<i>Training</i>	84.36%
	<i>Testing</i>	82.94%
AUC	<i>Training</i>	0.868%
	<i>Testing</i>	0.81%
Legends		

Discussion

This retrospective study was designed to create a tree to recognize the dietary risk factors for CAD. Decision tree is a data mining algorithm which is generally used for predicting medical conditions such as coronary artery disease (26). We observed that adjusted protein sits at the apex of the tree which indicated that high levels of protein intake were the most important risk factor for CAD. According to our tree, only protein intake could identify patients with coronary artery stenosis according to angiography from healthy participant up to 80%. Higher degrees of protein intake were associated with CAD. Dietary manganese was the second most important variable after protein. Interestingly, as shown in Table 2, 91.7% of those who had protein intake ≤ 80.076 g/day and Manganese > 2.305 mg/day were healthy while 94.18% of whom had protein > 80.076 g/day, Manganese ≤ 2.870 mg/day and biotin ≤ 36.376 $\mu\text{g/day}$ was categorized in CAD group. According to the results, the accuracy of the tree was 84.36% for training dataset and 82.94% for testing dataset, respectively.

There are a few studies available investigating the risk factors of CAD using data mining. Tayefi et al. carried out a data mining research in 2346 subjects by using a decision tree algorithm. They entered 10 variables including sex, age, triglyceride (TG), total cholesterol (TC), low density lipoprotein (LDL), high density lipoprotein (HDL), fasting blood glucose (FBG), high sensitivity C-reactive protein (hs-CRP), systolic blood pressure (SBP) and diastolic blood pressure (DBP) in the decision tree model. They concluded hs-CRP was the most important risk factor of CAD and they also found FBG, sex and age were other risk factors of CAD. They reported the accuracy of 95.3% for their tree (17). Moreover, Xing et al. evaluated the effect of some variables including, tumor necrosis factor- α (TNF- α), interleukin-6 (IL-6), interleukin-8 (IL-8), hs-CRP, methylputrescine oxidase-1 (MPO1), troponin I-2 (TNI2), sex, age, smoking,

hypertension, and diabetes on prediction of CAD survival using three algorithms including decision tree. They found that that decision tress models have accuracy of 89.6% (27).

To the best of our knowledge this is the only study using data mining algorithms for risk stratification of angiographic results considering dietary intake as potential factors. However, many studies have examined the effects of dietary intake on CAD prediction using other methodologies. Nazeminezhad et al, divided the population of study into three groups: 1) those with considerable disease (> 50% occlusion), 2) individuals with < 50% coronary artery occlusion, and 3) control group. After evaluating the dietary intake using a 24-h dietary recall method and dietary analysis, they found that those in control group have less dietary protein intake and higher manganese intake than that in the other two groups (2). In an 18 years follow-up study, in line with our findings, the researchers used a validated food-frequency questionnaire at 4 time points to assess nutrients intake. They observed that higher dietary vegetable protein significantly reduces risk of fatal ischemic heart disease. They also found that intake of animal protein is associated with ischemic heart disease occurrence in healthy men (28). Furthermore, it is previously shown that low consumption of protein and minerals (e.g. manganese) and high consumption of carbohydrate and fat is associated with having more severe CAD (29). Moreover, the data of a previous research on 36 adults showed that high dietary protein/ meat intake (more than 0.8 g protein/kg body weight/24 hours) induced CAD progression by increasing the lipid deposition, inflammation and coagulation pathways.

Regarding to the association between Manganese as a micro-nutrient and CAD, Manganese induces synthesis of cholesterol and fatty acids in liver (30). Manganese is also a part of enzymes superoxide dismutase and adenylyl cyclase enzymes involved in antioxidant mechanism (31, 32).

There are several possible explanations for this difference between studies. The methodology, the type of protein intake (of vegetable or animal) and questionnaire using for dietary intake are the most important factors responsible for this diversity.

Because of the increased CAD prevalence and consequently the heavy financial pressure on the society, finding ways to effectively predict this disease is a major desire of healthcare communities (33). Data mining might be used to notify individualized preventive actions and also define the impacts of each variables on the studied association. However, data mining has some limitations. It is a complicated method that needs specific knowledge and skills. In addition, each application created many rules and selection the meaningful ones requires experience.

Conclusion

Machine learning could be a powerful tool for risk stratification of diseases including CAD. Here we indicated considering dietary intake of protein and manganese along with zinc, biotin and cholesterol could predict CAD with accuracy of almost 85%.

Abbreviations

CAD Coronary artery disease

DBP Diastolic blood pressure

DT Decision tree

FFQ Food frequency questionnaire

HDL High density lipoprotein

Hs-CRP High sensitive-C-reactive protein

IL-6 Interleukin-6

IL-8 Interleukin-8

LDL Low density lipoprotein

MPO-1 Methylputrescine oxidase-1

QUEST Quick, Unbiased, Efficient Statistical Tree

SBP Systolic blood pressure

TC Total cholesterol

TG Triglyceride

TNF- α Tumor necrosis factor- α

TNI2 Troponin I-2

Declarations

Ethics approval and consent of participant:

The study protocol was given approval by the Ethics Committee of Mashhad University of Medical Sciences and written informed consent was obtained from participants.

Consent of publication:

Not applicable.

Availability of data and materials:

Not applicable.

Competing interests:

There is no competing interest.

Funding:

This study was supported by Mashhad University of Medical Sciences.

Author contributions

Study concept and design: Ghayour-Mobarhan, Esmaily, Moohebaty; data gathering: Tajfard; Analysis and interpretation of data: Saffar Soflaei, Shamsara; Drafting of the manuscript: Saffar Soflaei, Shamsara, Sahranavard, Shabani, Asadi; Critical revision of the manuscript for important intellectual content: Saffar Soflaei, Ghayour-Mobrahan, Ferns.

Acknowledgements:

We would like to thank Mashhad University of Medical Sciences for supporting this study.

References

References:

1. Mohammad AM, Jehangeer HI, Shaikhow SK. Prevalence and risk factors of premature coronary artery disease in patients undergoing coronary angiography in Kurdistan, Iraq. *BMC Cardiovasc Disord.* 2015;15(1):155.
2. Nazeminezhad R, Tajfard M, Latiff L, Mouhebaty M, Esmaily H, Ferns G, et al. Dietary intake of patients with angiographically defined coronary artery disease and that of healthy controls in Iran. *Eur J Clin Nutr.* 2014;68(1):109–13.
3. Abdar M. Using Decision Trees in Data Mining for Predicting Factors Influencing of Heart Disease. *Carpathian Journal of Electronic & Computer Engineering.* 2015;8(2).
4. Doris M, Newby DE. Coronary CT angiography as a diagnostic and prognostic tool: perspectives from the SCOT-HEART trial. *Curr Cardiol Rep.* 2016;18(2):18.
5. Mente A, de Koning L, Shannon HS, Anand SS. A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. *Arch Intern Med.* 2009;169(7):659–

69.

6. Talaei M, Koh WP, Yuan JM, van Dam RM. DASH Dietary Pattern, Mediation by Mineral Intakes, and the Risk of Coronary Artery Disease and Stroke Mortality. *Journal of the American Heart Association*. 2019;8(5):e011054.
7. Ford ES, Jiao J. Serum magnesium and ischaemic heart disease: findings from a national sample of US adults. 1999;28(4):645–51.
8. Singh RB, Niaz MA, Rastogi SS, Bajaj S, Gaoli Z, Shoumin, Z, Jot A, Co N. Current zinc intake and risk of diabetes and coronary artery disease and factors associated with insulin resistance in rural and urban populations of North India. 1998;17(6):564–70.
9. Aburto NJ, Hanson S, Gutierrez H, Hooper L, Elliott P, Cappuccio FP, B. Effect of increased potassium intake on cardiovascular risk factors and disease: systematic review and meta-analyses. 2013;346:f1378.
10. Fang X, Wang K, Han D, He X, Wei J, Zhao L, et al. Dietary magnesium intake and the risk of cardiovascular disease, type 2 diabetes, and all-cause mortality: a dose–response meta-analysis of prospective cohort studies. 2016;14(1):210.
11. Aburto NJ, Ziolkovska A, Hooper L, Elliott P, Cappuccio FP, Meerpohl J, J, B. Effect of lower sodium intake on health: systematic review and meta-analyses. 2013;346:f1326.
12. Yusuf S, Sleight P, Pogue J, Bosch J, Davies R, Dagenais G, J, N, E, J, M. The HOPE Investigators: effects of an angiotensin converting enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. 2000;342:145–53.
13. Lancet HP, S, C, G, J, T. MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20 536 high-risk individuals: a randomised placebo-controlled trial. 2002;360(9326):23–33.
14. Posner BM, Cobb JL, Belanger AJ, Cupples LA, D'Agostino RB, Stokes J, J, A, O, I, M. Dietary lipid predictors of coronary heart disease in men: the Framingham Study. 1991;151(6):1181–7.
15. Preis SR, Stampfer MJ, Spiegelman D, Willett WC, Rimm, E, B, J, T, A, j, o, c, n. Lack of association between dietary protein intake and risk of stroke among middle-aged men. 2009;91(1):39–45.
16. Pedersen AN, Kondrup J, Børshøj E, J, F, research n. Health effects of protein intake in healthy adults: a systematic literature review. 2013;57(1):21245.
17. Tayefi M, Tajfard M, Saffar S, Hanachi P, Amirabadizadeh AR, Esmaeily H, et al. hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm. *Comput Methods Programs Biomed*. 2017;141:105–9.
18. Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. *Value in Health*. 2019;22(7):808–15.
19. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *European heart journal*. 2017;38(7):500–7.

20. Groselj C, Kukar M, Fettich J, Kononenko I, editors. Machine learning improves the accuracy of coronary artery disease diagnostic methods. *Computers in Cardiology 1997; 1997: IEEE.*
21. Joloudari JH, Hassannataj Joloudari E, Saadatfar H, GhasemiGol M, Razavi SM, Mosavi A, et al. Coronary artery disease diagnosis; ranking the significant features using a random trees model. *Int J Environ Res Public Health.* 2020;17(3):731.
22. Rokach L, Maimon OZ. *Data mining with decision trees: theory and applications: World scientific; 2008.*
23. Lin C-L, Fan C-L. Evaluation of CART, CHAID, and QUEST algorithms: a case study of construction defects in Taiwan. *Journal of Asian Architecture Building Engineering.* 2019;18(6):539–53.
24. Hu FB, Stampfer MJ, Rimm E, Ascherio A, Rosner BA, Spiegelman D, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. *Am J Epidemiol.* 1999;149(6):531–40.
25. Mackerras D. Energy adjustment: the concepts underlying the debate. *J Clin Epidemiol.* 1996;49(9):957–62.
26. Patel AKSSB, Shukla D. A data mining technique for prediction of coronary heart disease using neuro-fuzzy integrated approach two level. *International Journal of Engineering Computer Science.* 2013;2:09.
27. Xing Y, Wang J, Zhao Z, editors. Combination data mining methods with new medical data to predicting outcome of coronary heart disease. *2007 International Conference on Convergence Information Technology (ICCIT 2007); 2007: IEEE.*
28. Preis SR, Stampfer MJ, Spiegelman D, Willett WC, Rimm EB. Dietary protein and risk of ischemic heart disease in middle-aged men. *Am J Clin Nutr.* 2010;92(5):1265–72.
29. Mahalle N, Garg M, Naik S, Kulkarni M. Association of dietary factors with severity of coronary artery disease. *Clinical nutrition ESPEN.* 2016;15:75–9.
30. Delimaris I. Adverse effects associated with protein intake above the recommended dietary allowance for adults. *International Scholarly Research Notices.* 2013;2013.
31. Moustafa SR. Biochemical Estimation of Trace Elements Manganese (Mn), Cobalt (Co), Calcium (Ca) and Iron (Fe) in Patients with Cardiovascular Diseases. *Tikrit Journal of Pure Science.* 2011;16(4):118–23.
32. Fujimoto H, Taguchi J-i, Imai Y, Ayabe S, Hashimoto H, Kobayashi H, et al. Manganese superoxide dismutase polymorphism affects the oxidized low-density lipoprotein-induced apoptosis of macrophages and coronary artery disease. *European heart journal.* 2008;29(10):1267–74.
33. Ayatollahi H, Gholamhosseini L, Salehi M. Predicting coronary artery disease: a comparison between two data mining algorithms. *BMC Public Health.* 2019;19(1):1–9.

Figures

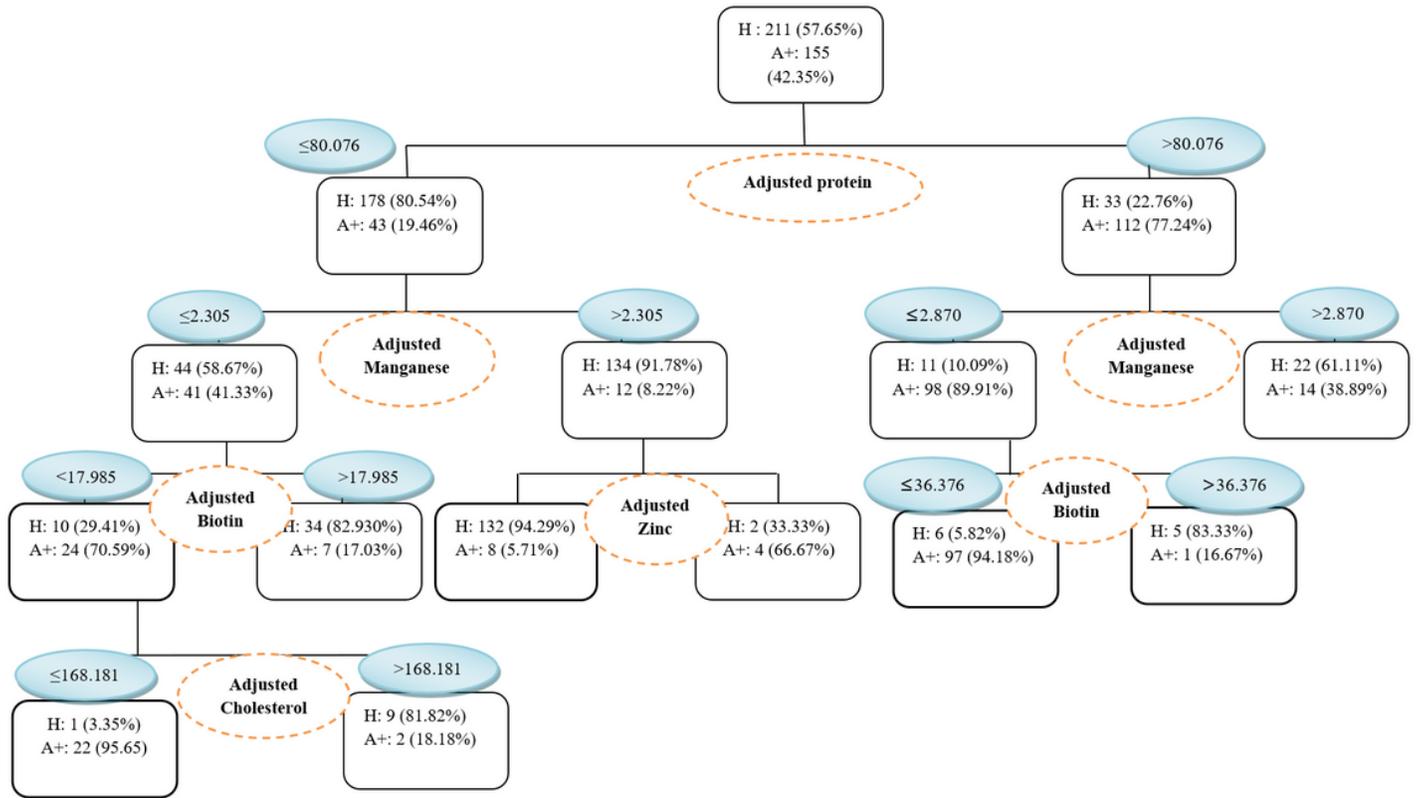


Figure 1

Final decision tree with 12 leaves and 4 layers