Molecular Biology Reports
https://doi.org/10.1007/s11033-018-4236-2

ORIGINAL ARTICLE

Machine learning algorithm-based risk prediction model of coronary artery disease

Shaik Mohammad Naushad[1,5] · Tajamul Hussain[2] · Bobbala Indumathi[3] · Khatoon Samreen[1] · Salman A. Alrokayan[4] · Vijay Kumar Kutala[3]

Received: 21 April 2018 / Accepted: 2 July 2018

Abstract

In view of high mortality associated with coronary artery disease (CAD), development of an early predicting tool will be beneficial in reducing the burden of the disease. The database comprising demographic, conventional, folate/xenobiotic genetic risk factors of 648 subjects (364 cases of CAD and 284 healthy controls) was used as the basis to develop CAD risk and percentage stenosis prediction models using ensemble machine learning algorithms (EMLA), multifactor dimensionality reduction (MDR) and recursive partitioning (RP). The EMLA model showed better performance than other models in disease (89.3%) and stenosis prediction (82.5%). This model depicted hypertension and alcohol intake as the key predictors of CAD risk followed by cSHMT C1420T, GCPII C1561T, diabetes, GSTT1, CYP1A1 m2, TYMs 5′-UTR 28 bp tandem repeat and MTRR A66G. MDR and RP models are in agreement in projecting increasing age, hypertension and cSHMTC1420T as the key determinants interacting in modulating CAD risk. Receiver operating characteristic curves exhibited clinical utility of the developed models in the following order: EMLA (C = 0.96) > RP (C = 0.83) > MDR (C = 0.80). The stenosis prediction model showed that xenobiotic pathway genetic variants i.e. CYP1A1 m2 and GSTT1 are the key determinants of percentage of stenosis. Diabetes, diet, alcohol intake, hypertension and MTRR A66G are the other determinants of stenosis. These eleven variables contribute towards 82.5% stenosis. To conclude, the EMLA model exhibited higher predictability both in terms of disease prediction and stenosis prediction. This can be attributed to higher number of iterations in EMLA model that can increase the prediction accuracy.

Keywords Coronary artery disease · Folate and xenobiotic pathways · Ensemble machine learning algorithm · Multifactor dimensionality reduction · Recursive partitioning

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11033-018-4236-2) contains supplementary material, which is available to authorized users.

✉ Shaik Mohammad Naushad
naushad@sandor.co.in; naushadsm@gmail.com
https://www.researchgate.net/profile/Naushad_Shaik_Mohammad/contributions

Tajamul Hussain
thussain@ksu.edu.sa
https://www.researchgate.net/profile/Tajamul_Hussain2/contributions

Bobbala Indumathi
indubobbala@gmail.com

Khatoon Samreen
samreen@sandor.co.in

Salman A. Alrokayan
salrokayan@ksu.edu.sa
https://www.researchgate.net/profile/Salman_Alrokayan/contributions

Vijay Kumar Kutala
vijaykutala@gmail.com
https://www.researchgate.net/profile/Vijay_Kumar_Kutala/contributions

[1] Sandor Lifesciences Pvt Ltd, Banjara Hills, Road No. 3, Hyderabad, India

[2] Center of Excellence in Biotechnology Research, College of Science, King Saud University, Riyadh, Saudi Arabia

[3] Department of Clinical Pharmacology and Therapeutics, Nizam's Institute of Medical Sciences, Hyderabad, India

[4] Department of Biochemistry, College of Science, King Saud University, Riyadh, Saudi Arabia

[5] Biochemical Genetics & Pharmacogenomics, Sandor Lifesciences Pvt Ltd, Banjara Hills, Road No. 3, Hyderabad 500034, India

Published online: 11 July 2018

Introduction

Coronary artery disease (CAD) is one of the major leading causes of death all over the world with an estimated 17.7 million deaths from CAD in 2015. Over three quarters of these deaths are taking place in low- and middle-income countries. This prompts that subjects with conventional risk factors such as diabetes, hypertension, dyslipidemia or family history of CAD should be tested for early atherosclerotic changes to facilitate early detection and management of the disease [1]. Among the nutritional factors that are likely to modulate risk for CAD, the most prominent being folate and B-complex vitamins whose deficiencies induce oxidative stress by increasing homocysteine levels and by interfering with phase II detoxification of polycyclic aromatic hydrocarbons [2].

Lower *S*-adenosyl methionine and 5-methyltetrahydrofolate levels, and higher total plasma homocysteine levels were observed in CAD patients [3]. ER alpha hypermethylation was reported in coronary atherosclerotic plaques when compared to normal aorta [4]. Treatment with 3-deazaadenosine was shown to prevent smooth muscle cell proliferation and neointima formation by interfering with Ras methylation [5]. LDL was found to induce expression of p66shc via hypomethylation of its promoter thus mediating a dysfunctional endothelial cell surface with proadhesive and procoagulant features [6].

Methylation of ATP-binding cassette A1 (ABCA1) was shown to lower HDL levels thus increasing risk for CAD [7]. L5, the most negatively charged sub-fraction of low-density lipoprotein that is capable of inducing apoptosis, was shown to inhibit fibroblast growth factor-2 (FGF-2) by inducing hypermethylation of its promoter [8]. Hyperlipidemia was shown to be associated with methylation of ATP-binding cassette, sub-family G (WHITE), member 1 (ABCG1), lipase, hepatic (LIPC) and phospholipid transfer protein (PLTP) [9]. Hypermethylation of dimethylarginine dimethylaminohydrolase 2 (DDAH2) was shown to impair function of endothelial progenitor cells thus playing an important role in the pathophysiology of CAD [10].

Further, differentially methylated regions in folate pathway genetic loci i.e. transcobalamin-2 (TCN2) promoter, cysthathionine beta synthase (CBS) 5′UTR, aminomethyltransferase (AMT), paraoxonase/arylesterase 1 (PON1) were observed [11]. The aberrations in folate metabolism were shown to induce altered gene expression of extracellular superoxide dismutase (EC-SOD), glutathione-*S*-transferase (GST)P1, and BCL2/Adenovirus E1B 19 kDa protein-interacting protein 3 (BNIP3) thus contributing to the increased oxidative stress and increased susceptibility to CAD [12].

Aim

These studies are point towards the interplay among the conventional risk factors, folate and xenobiotic metabolic pathways and epigenome modulating risk for CAD. In order to explore these interactions, we have used the data from two of our earlier studies on folate [13] and xenobiotic [14] metabolic pathways for the development of coronary artery stenosis prediction model. The results of the first study showed that methylene tetrahydrofolate reductase (MTHFR) C677T and methyltetrahydrofolate homocysteine methyltransferase reductase (MTRR) A66G increase CAD risk by 1.61 and 1.92-folds, while thymidylate synthase (TYMS) 5′-UTR 2R-allele reduces the CAD risk by 34% [13]. The results of the second study showed that two cytochrome P450 1A1 (CYP1A1) haplotypes (m1–m2–m3) i.e. CAC and TGC increase the CAD risk, while TAC haplotype confers protection [14]. Glutathione-*S*-transferase (GSTT1) null genotype was shown to increase CAD risk [14]. In the current study, a CAD risk prediction was developed using demographic (n = 5), conventional (n = 3) and genetic (n = 14) variables as input variables and percentage of stenosis as output to explore the potential gene–gene and gene-environmental interactions. To achieve this objective, multifactor dimensionality reduction (MDR), recursive partitioning (RP) and machine learning algorithms (EMLA) were employed.

Methods

Study population

The study population comprised of 648 subjects (364 cases with documented CAD and 284 healthy controls). Cases were angiographically documented to have CAD with the stenosis in the range of 50–90% (67.34 ± 13.00%) and controls were ethnicity-matched to cases. The subjects with chronic inflammatory disease, immunological disease and cancer were excluded.

Demographic data such as age, gender, body mass index (BMI), smoking and alcohol intake were obtained from the subjects. Along with the demographic data, conventional risk-factors like hypertension, diabetes, hyperlipidemia were also included. The study protocol was approved by the Institutional ethical committee of Nizam's Institute of Medical Sciences, Hyderabad. All the subjects consented for the study.

Analysis of genetic polymorphisms

Whole blood samples collected in EDTA were used for genomic DNA extraction using phenol–chloroform extraction method. PCR-RFLP method was used for the analysis of glutamate carboxypeptidase II (GCPII) C1561T, reduced folate carrier 1 (RFC1) G80A, cytosolic serine hydroxyl methyltransferase (cSHMT) C1420T, MTHFR C677T, methionine synthase (MTR) A2756G, MTRR A66G, CYP1A1 m1, CYP1A1 m2, CYP1A1 m3, catecholamine-*O*-methyl transferase (COMT) H108L polymorphisms. PCR-AFLP method was used for the analysis of TYMS 5′-UTR 28 bp tandem repeat. Multiplex PCR method was used to analyze deletions in GSTT1 and GSTM1.

Risk prediction models

Machine learning algorithm

We have developed Ensemble machine learning algorithms using the risk factors for CAD as the input variables and CAD risk or percentage of stenosis as the predictor. Ensemble constructs a set of classifiers and classifies new data points by taking a note of their predictions. Bayesian averaging or error-correcting output coding, bagging and boosting were used as the basis of the model.

Multifactor dimensionality reduction (MDR)

We have developed MDR model by incorporating the 22 variables as x1, x2, x3, …, x22 and output variable as class. Being a non-parametric and model-free method, MDR reduced the dimensionality of multi-locus information and identified most significant variables associated with an increased risk for CAD in terms of univariate, bivariate and trivariate analysis along with providing Frutcherman-Rheingold plots that depict strength of association in terms of entropy. This analysis was carried out using the computational website http://www.multifactordimensionalityreduction.org.

Recursive partitioning

We have constructed a regression tree using 22 variables as input variables and categorical variable as output variable. The stem of the tree depicts the main predictor while it branches show various interactions and decision levels. This analysis was carried out using the computational website http://www.wessa.net.

Clinical utility of the risk prediction models

We have used receiver-operating characteristic (ROC) curves as the indices of clinical utility of these models as they depict overall accuracy in terms of area under the curve and represent true positive rate (sensitivity) vs. false positive rate (1 − specificity). This analysis was carried out using GraphPad Prism software.

Results

Univariate analysis

Univariate analysis of demographic and genetic variables was tabulated as Table 1. This analysis revealed that vegetarian diet, alcohol intake, diabetes and hypertension as nutritional or life style risk factors. Among the folate pathway genetic variants, GCPII C1561T, MTHFR C677T and MTRR A66G are risk factors for CAD while TYMS 5′-UTR 28 bp tandem repeat and cSHMT C1420T confer protection against CAD. Among the xenobiotic pathway

Table 1 Demographic and genetic characteristics of cases and controls

Variable	Cases	Controls	P value
Age	53.8 ± 10.9	52.5 ± 9.8	0.12
Gender	293:71	225:56	0.97
Body mass index (kg/m^2)	24.1 ± 3.4	24.2 ± 3.5	0.80
Non-vegetarian diet	215:70	240:37	0.001*
Smoking	84:180	82:194	0.66
Alcohol intake	70:217	30:246	< 0.0001*
Diabetes	95:201	18:258	< 0.0001*
Hypertension	127:163	33:242	< 0.0001*
GCPII C1561T	12.45%	5.12%	< 0.0001*
RFC1G80A	42.56%	39.73%	0.32
cSHMT C1420T	44.44%	58.54%	0.001*
TYMS 5′-UTR 28 bp tandem repeat	26.95%	35.84%	< 0.0001*
MTHFR C677T	14.33%	8.75%	0.005*
MTR A2756G	30.05%	29.57%	0.18
MTRR A66G	36.68%	28.91%	0.006*
CYP1A1 m1	22.32%	17.02%	0.09
CYP1A1 m2	17.66%	12.99%	0.04*
CYP1A1 m4	0.57%	0.36%	0.93
COMT H108L	17.71%	19.32%	0.38
GSTT1	11.47%	6.91%	0.01*
GSTM1	9.69%	9.57%	1.00

Continuous variables were presented in mean ± standard deviation format while categorical variables were presented either in raw numbers or in percentage. Student *t* test and Fisher exact test were performed for continuous and categorical variables, respectively

genetic variants, CYP1A1 m1 and GSTT1 null variants are associated with CAD risk.

CAD risk prediction using machine learning algorithm (EMLA)

As shown in Fig. 1, EMLA model based on classification template was developed using 11 risk factors identified in univariate analysis as input variables while presence or

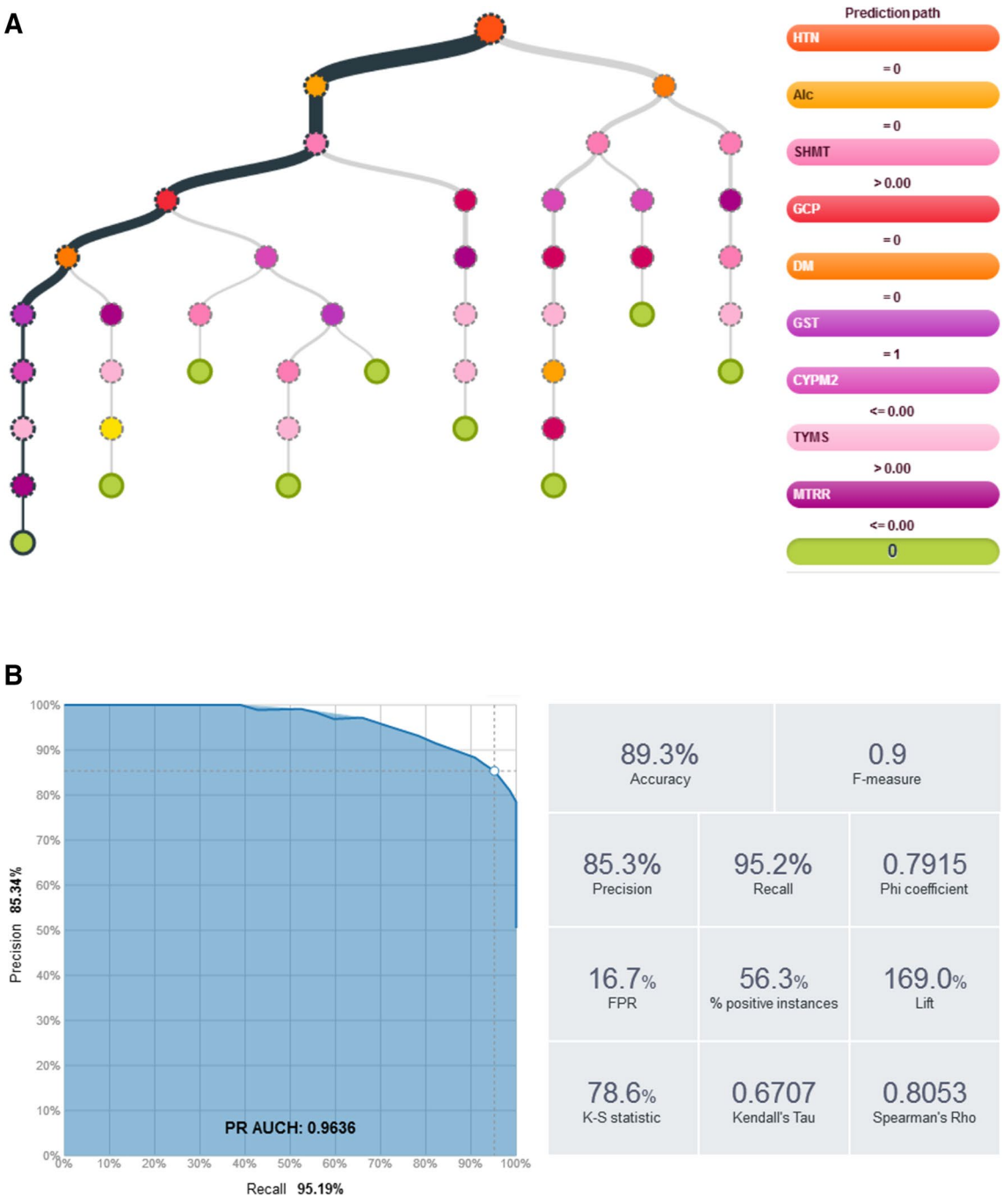

Fig. 1 The Ensemble machine learning-based prediction model. **a** Architecture of machine learning algorithm: machine learning algorithm identified hypertension, alcohol intake, cSHMT C1420T, GCPII C1561T, diabetes mellitus, GSTT1, CYP1A1 m2, TYMS 5′-UTR 28 bp tandem repeat and MTRR A66G as the determinants of CAD risk. **b** Receiver operating characteristic (ROC) curve depicting the clinical utility of this model: The area under the ROC curve was 0.96 with 89.3% accuracy in predicting CAD risk

absence of CAD was used as the output variable. This risk prediction model had 89.3% accuracy in predicting CAD risk with a precision of 85.3% and recall rate of 95.2%. True positive and false positive rates were 95.19 and 16.67%. The area under the ROC curve was 0.96. With the given true positive and true negative predictions with type I alpha error of 0.05 for 80% of power, required sample size is 345. We have studied 648 subjects in total, substantiating the statistical power of the study.

The life style risk factors namely hypertension, alcohol intake and diabetes are in the apex of the model as key determinants. Among the genetic risk factors, cSHMT C1420T was the key determinant of CAD risk.

The simulations of the developed model revealed that vegetarian or non-vegetarian diet in the absence of other risk factors is not associated with CAD risk. Alcohol intake was shown to exert CAD risk, which is negated in the presence of cSHMT 1420 TT-genotype. However, presence of variant alleles at MTHFR and MTRR loci synergistically counteract the protection conferred by cSHMT 1420 TT-genotype.

Gender-specific associations

EMLA model of men showed age as an important contributor in modulating CAD risk. In subjects with age > 52 years, cSHMT CC-genotype is a risk factor for CAD. In diabetic subjects with age > 52 years, low dietary folate intake and TYMS 5′-UTR 28 bp tandem repeat are risk factors for CAD. In non-diabetic subjects with age > 52 years, BMI > 26.88 kg/m^2 and low dietary folate intake are risk factors for CAD. In subjects with age < 52 years, presence of hypertension and BMI > 23.8 kg/m^2 were associated with CAD risk. In non-hypertensive subjects with age < 52 years, alcohol consumption followed by presence of CYP1A1 m2 or cSHMT CC and CT genotypes is associated with CAD risk (Supplementary Fig. 1).

In post-menopausal women, diabetes is a risk factor for CAD. In the absence of diabetes, cSHMT CC, CYP1A1 m2 and TYMS 5′-UTR 3R3R are the risk factors for CAD. Among these women with low folate intake, MTHFR C677T is a risk factor for CAD. In pre-menopausal women, family history of CAD is a risk factor for CAD. Diabetes and low folate intake are associated with CAD risk. Pre-menopausal status, absence of diabetes, lower BMI and CYP1A1 m2 genotype are associated with reduced risk for CAD in women. MTHFR C677T was shown to be a risk factor for CAD in post-menopausal women with low folate intake (Supplementary Fig. 2).

Multifactor dimensionality reduction

As shown in Fig. 2a, hypertension is the most significant risk factor for CAD. As shown in Fig. 2b, alcohol intake is another important risk factor for CAD. Synergistic risk inflation was observed in hypertensive subjects with alcohol intake. As shown in Fig. 2c, subjects with cSHMT TT-genotype are protected against alcohol-mediated CAD risk. As shown in Fig. 3, Frutcherman-Rheingold plot depicted the following order of strength: hypertension > cSHMT C1420T > alcohol intake. Synergistic interaction between hypertension and alcohol intake increases CAD risk. The cSHMT C1420T was shown to have counteracting interactions alcohol intake and hypertension in reducing risk for CAD. The area under the curve for MDR model was 0.80 (95% CI 0.75–0.85), p < 0.0001.

Recursive partitioning

As shown in Fig. 4, age in decades is the most important predictor in this model. In subjects with age > 50 years, cSHMT 1420 CC genotype is associated with CAD risk. In subjects with age < 50 years, lack of hypertension and no alcohol consumption reduces the CAD risk significantly. The area under the curve of ROC was 0.83 (95% CI 0.79–0.88), p < 0.0001.

Stenosis prediction using machine learning algorithm

A stenosis prediction model was developed using machine learning tools (Fig. 5a). This model highlighted the importance of xenobiotic pathway genetic variants i.e. CYP1A1 m2 and GSTT1 in modulating percentage of stenosis. Among the life style risk factors, diabetes, diet, alcohol and hypertension were found to be key determinants of percentage of stenosis. Among the folate pathway genetic variants, MTRR A66G alone was shown to be one of the determinants of percentage of stenosis. These eleven variables contribute towards 82.5% stenosis (Fig. 5b).

Discussion

In the current study, EMLA, MDR and RP based risk prediction models for CAD and EMLA based prediction models of percentage of stenosis were developed. The EMLA model being complex performed well both in predicting CAD risk and percentage of stenosis. Hypertension and alcohol intake are the key predictors of CAD risk in EMLA model followed by cSHMT C1420T, GCPII C1561T, diabetes, GSTT1, CYP1A1 m2, TYMs 5′-UTR 28 bp tandem repeat and MTRR A66G. The MDR and RP models are in agreement with each other in portraying hypertension, cSHMT C1420T and alcohol intake as the key determinants of CAD risk out of the 22 variables tested. In subjects with age < 50 years, hypertension and alcohol intake were observed as

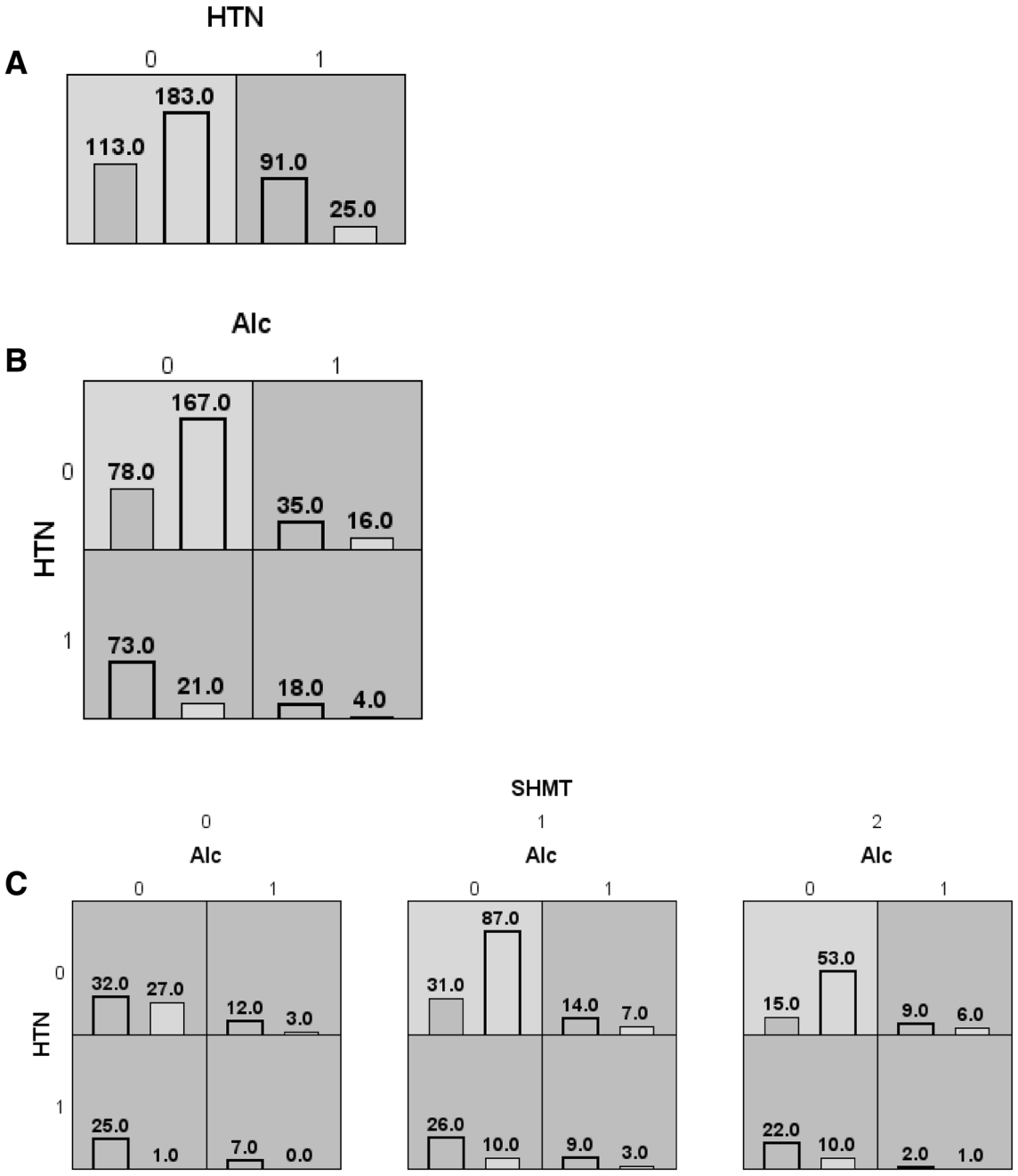

Fig. 2 Multifactor dimensionality reduction analysis. **a** Univariate analysis: hypertension is the key determinant of CAD risk. **b** Bivariate analysis: reduced risk for CAD was observed in non-hypertensive and non-alcoholic subjects. Alcohol intake and hypertension are associated individually and synergistically with increased risk for CAD. **c** Trivariate analysis: cytosolic SHMT C1420T confers protection against alcohol intake-mediated CAD risk. The light and dark background suggests no risk and high risk for CAD, respectively. Two bars in each block represent number of cases and controls

key modulators of CAD risk. Gender-based differences in association with CAD were also explored. Diabetes is a risk factor for CAD specifically in men with age > 52 years and in both pre- and post-menopausal women. In men with age < 52 years, hypertension is the main risk factor for CAD. In both genders, obesity is an important risk factor for CAD. Alcohol consumption followed by presence of cSHMT and CYP1A1 m2 variant alleles was shown to exert CAD risk in non-hypertensive men. Even in post-menopausal women cSHMT and CYP1A1 m2 are risk factors for CAD. In pre-menopausal women, family history of CAD, diabetes and low folate intake increase the risk for CAD.

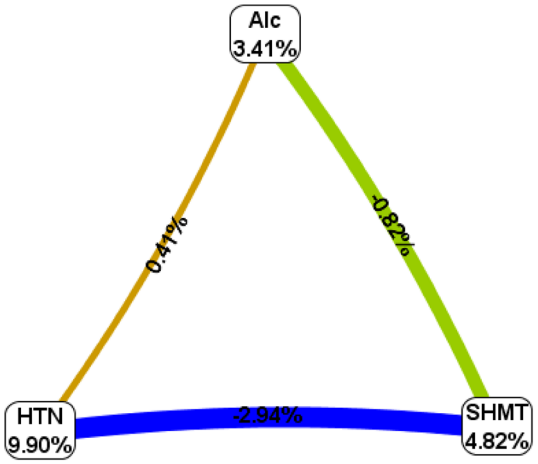

Fig. 3 Frutcherman-Rheingold plot depicting interactions between variables. This illustrates the strength of bonding based on entropy levels. The order of strength of association of variables: hypertension (HTN) > cSHMT C1420T > alcohol intake. Synergistic interaction (entropy: 0.41%) was observed between hypertension and alcohol while cSHMT C1420T counteracts the risk associated with alcohol intake (entropy: −0.82) and hypertension (entropy: −2.94%)

The developed EMLA model showed the association of alcohol intake with CAD risk followed by counteracting effect of cSHMT 1420 TT-genotype in negating this risk. It was reported that alcohol intake is associated with erythrocyte folate deficiency [15] and presence of cSHMT 1420 TT-genotype increases H-bonding interactions between cSHMT protein and tetrahydrofolate thus increasing the folate flux through induction of futile folate cycle [16]. However, the presence of variant alleles at MTHFR and MTRR loci synergistically impair remethylation of homocysteine significantly and contribute towards increased CAD risk [13]. The presence of 677C > T variant in MTHFR induces thermolability thus enhancing the propensity of MTHFR active dimer into inactive monomers with decreased FAD-binding capacity [17]. In the presence of low dietary folate intake along with MTHFR 677 C > T, the 5-methyl tetrahydrofolate synthesis is significantly impaired thus affecting *S*-adenosylmethionine levels. *S*-adenosylmethionine is essential for catecholamine-*O*-methyl transferase activity necessary for the conversion of catechol estrogens to methoxy estrogens. In post-menopausal women with low folate and MTHFR 677 C > T this conversion might be affected due to low SAM thus increasing catechol estrogen leading to formation of quinones and semiquinones that induced oxidative lesions in DNA. Consistent with our observation, Tanis et al. reported twofold increase in risk for myocardial infarction in women with MTHFR 677 TT-genotype when their folate levels are below the median [18].

CYP1A1 m2, GSTT1, diabetes, diet, alcohol intake, hypertension and MTRR A66G are the key determinants of stenosis. Our results are in agreement with another risk prediction model based on ANN which demonstrated hypertension, diabetes, hyperlipidemia and homocysteine as the determinants of coronary artery stenosis [19]. Keeping diabetes and hypertension under control followed by regular exercise can drastically reduce the percentage of stenosis as depicted by this model. This is further supported by a recent study which depicted positive correlation between the admission HbA1C level and number of affected vessels

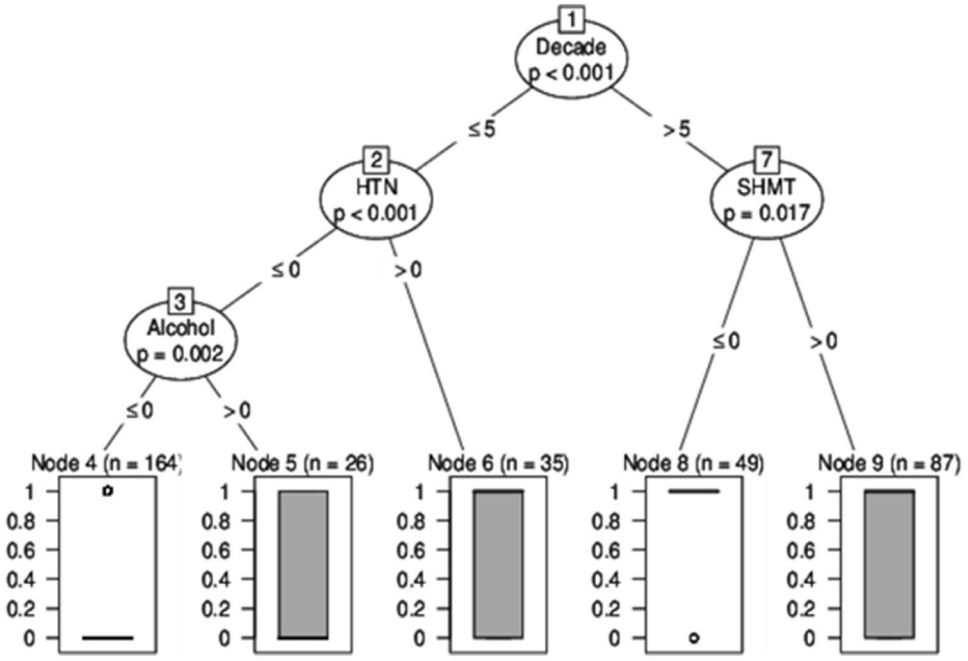

Fig. 4 Recursive partitioning model for CAD prediction. This illustrates age in decade as the main predictor followed by cSHMT in subjects with > 50 years and hypertension in subjects < 50 years. Subjects with < 50 years are at lower risk of CAD is they are non-alcoholic and non-hypertensive

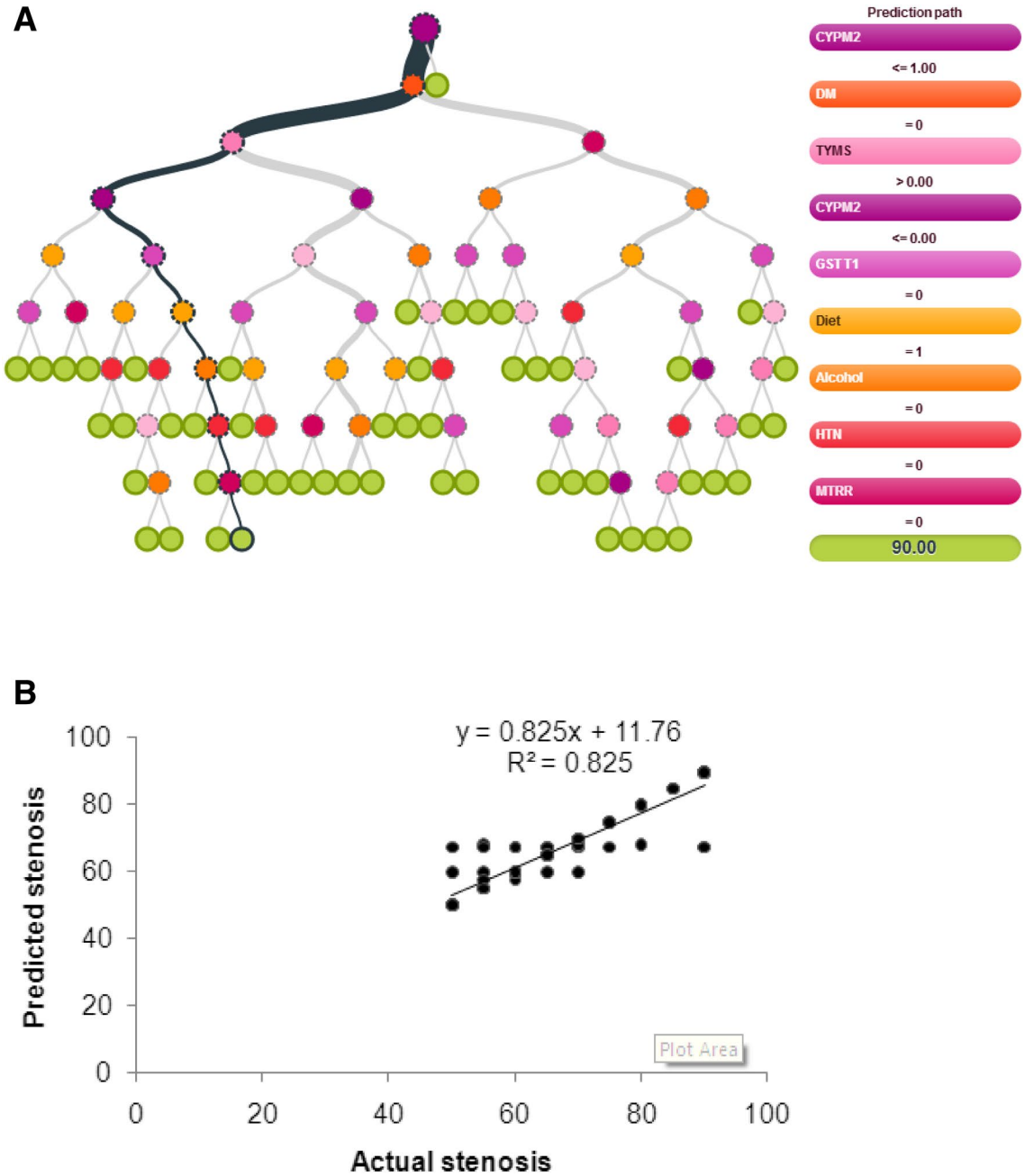

Fig. 5 Risk prediction model of stenosis. **a** The ensemble machine learning algorithm: This projects CYP1A1 m2, diabetes mellitus, TYMS 5′-UTR 28 bp tandem repeat, GSTT1, diet, alcohol intake, hypertension, MTRR A66G as the predictors of stenosis. **b** Correlation between actual stenosis and predicted stenosis: machine learning algorithm explains 82.5% variability in stenosis

[20]. CYP1A1 and GSTT1 polymorphisms were shown to increase oxidative DNA damage and thus contribute to CAD risk [14].

A recent large scale study has developed a similar CAD risk prediction model using age, gender, type of chest pain, diabetes, hypertension, dyslipidemia, smoking status and laboratory data as the predictors [21]. This model showed an AUC of 0.72. The predictability of our model is higher due to the presence of additional genetic variables. The risk reduction by a decade in subjects with cSHMT TT genotype can be attributed to decreased oxidative stress due to increased folate pool and antioxidant status as reported earlier [22]. Carotid intima-media thickness was shown to have positive association with CBS rs2851391 × MTR rs180508 while exhibiting inverse association with vitamin B_6, B_{12} and folate [23], which is consistent with the current study in

highlighting the role of folate pathway aberrations in CAD risk prediction.

The inverse association of regular exercise with percentage of stenosis is consistent with a recent study where in exercise-based cardiac rehabilitation improved the peak oxygen intake with better quality of life [24].

The major strengths of the current study are application of three different risk prediction models i.e. EMLA, MDR and RP. The agreement between MDR and RP models in exploring interactions among variables made the associations more robust. The limitations of the current study were lack of data on HbA1c, C-reactive protein and micronutrient status of CAD patients. Further, the large studies are warranted by incorporating these variables in the studied models to increase the precision in risk prediction.

Conclusion

EMLA model outperformed MDR and RP models in predicting CAD risk. MDR and RP models are in agreement in depicting age in decades, hypertension and cSHMT C1420T as the key determinants that modulate age of onset of CAD. In younger subjects, controlling the blood pressure and avoiding alcohol intake was shown to reduce the risk for CAD. The demographic, conventional and folate/xenobiotic genetic risk factors together was found to predict the CAD risk with 89.3% accuracy. However, their contribution in explaining variability of percentage stenosis was 82.5%.

Acknowledgements The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research Group No. (RG-1435-066).

Compliance with Ethical Standards

Conflict of interest None of the authors has a financial or other relationship with other people or organizations that may inappropriately influence this work.

Ethical approval The study protocol was approved by Institutional Ethical committee of Nizam's Institute of Medical Sciences, Hyderabad. This study complied with the ethical principles outlined in the Declaration of Helsinki.

Informed consent Informed consent was obtained from all the study participants (cases as well as controls) during their enrollment for the study.

References

1. Adam AM, Rehan A, Waseem N, Iqbal U, Saleem H, Ali MA, Shaikh AT, Godil A (2017) Prevalence of conventional risk factors and evaluation of baseline indices among young and elderly patients with coronary artery disease. J Clin Diagn Res 11(7):OC34–OC39
2. Mager A, Orvin K, Koren-Morag N, Lev IE, Assali A, Kornowski R, Shohat M, Battler A, Hasdai D (2009) Impact of homocysteine-lowering vitamin therapy on long-term outcome of patients with coronary artery disease. Am J Cardiol 104(6):745–749
3. Loehrer FM, Angst CP, Haefeli WE, Jordan PP, Ritz R, Fowler B (1996) Low whole-blood S-adenosylmethionine and correlation between 5-methyltetrahydrofolate and homocysteine in coronary artery disease. Arterioscler Thromb Vasc Biol 16(6):727–733
4. Post WS, Goldschmidt-Clermont PJ, Wilhide CC, Heldman AW, Sussman MS, Ouyang P, Milliken EE, Issa JP (1999) Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. Cardiovasc Res 43(4):985–991
5. Sedding DG, Tröbs M, Reich F, Walker G, Fink L, Haberbosch W, Rau W, Tillmanns H, Preissner KT, Bohle RM, Langheinrich AC (2009) 3-Deazaadenosine prevents smooth muscle cell proliferation and neointima formation by interfering with Ras signaling. Circ Res 104(10):1192–1200
6. Kim YR, Kim CS, Naqvi A, Kumar A, Kumar S, Hoffman TA, Irani K (2012) Epigenetic upregulation of p66shc mediates low-density lipoprotein cholesterol-induced endothelial cell dysfunction. Am J Physiol Heart Circ Physiol 303(2):H189–H196
7. Guay SP, Brisson D, Munger J, Lamarche B, Gaudet D, Bouchard L (2012) ABCA1 gene promoter DNA methylation is associated with HDL particle profile and coronary artery disease in familial hypercholesterolemia. Epigenetics 7(5):464–472
8. Chang PY, Chen YJ, Chang FH, Lu J, Huang WH, Yang TC, Lee YT, Chang SF, Lu SC, Chen CH (2013) Aspirin protects human coronary artery endothelial cells against atherogenic electronegative LDL via an epigenetic mechanism: a novel cytoprotective role of aspirin in acute myocardial infarction. Cardiovasc Res 99(1):137–145
9. Guay SP, Brisson D, Lamarche B, Gaudet D, Bouchard L (2014) Epipolymorphisms within lipoprotein genes contribute independently to plasma lipid levels in familial hypercholesterolemia. Epigenetics 9(5):718–729
10. Niu PP, Cao Y, Gong T, Guo JH, Zhang BK, Jia SJ (2014) Hypermethylation of DDAH2 promoter contributes to the dysfunction of endothelial progenitor cells in coronary artery disease patients. J Transl Med 12:170
11. Fiorito G, Guarrera S, Valle C, Ricceri F, Russo A, Grioni S, Mattiello A, Di Gaetano C, Rosa F, Modica F, Iacoviello L, Frasca G, Tumino R, Krogh V, Panico S, Vineis P, Sacerdote C, Matullo G (2014) B-vitamins intake, DNA-methylation of one carbon metabolism and homocysteine pathway genes and myocardial infarction risk: the EPICOR study. Nutr Metab Cardiovasc Dis 24(5):483–488
12. Lakshmi SV, Naushad SM, Reddy CA, Saumya K, Rao DS, Kotamraju S, Kutala VK (2013) Oxidative stress in coronary artery disease: epigenetic perspective. Mol Cell Biochem 374(1–2):203–211
13. Vijaya Lakshmi SV, Naushad SM, Rupasree Y, Seshagiri Rao D, Kutala VK (2011) Interactions of 5′-UTR thymidylate synthase polymorphism with 677C → T methylene tetrahydrofolate reductase and 66A → G methyltetrahydrofolate homocysteine methyltransferase reductase polymorphisms determine susceptibility to coronary artery disease. J Atheroscler Thromb 18(1):56–64
14. Lakshmi SV, Naushad SM, Saumya K, Rao DS, Kutala VK (2012) Role of CYP1A1 haplotypes in modulating susceptibility to coronary artery disease. Indian J Biochem Biophys 49(5):349–355
15. Sanvisens A, Zuluaga P, Pineda M, Fuster D, Bolao F, Juncà J, Tor J, Muga R (2017) Folate deficiency in patients seeking treatment of alcohol use disorder. Drug Alcohol Depend 180:417–422

16. Naushad SM, Radha Rama Devi A, Nivetha S, Lakshmitha G, Stanley AB, Hussain T, Kutala VK (2017) Neurofuzzy model of homocysteine metabolism. J Genet 96(6):919–926
17. Yamada K, Chen Z, Rozen R, Matthews RG (2001) Effects of common polymorphisms on the properties of recombinant human methylenetetrahydrofolate reductase. Proc Natl Acad Sci USA 98(26):14853–14858
18. Tanis BC, Blom HJ, Bloemenkamp DG, van den Bosch MA, Algra A, van der Graaf Y, Rosendaal FR (2004) Folate, homocysteine levels, methylenetetrahydrofolate reductase (MTHFR) 677C --> T variant, and the risk of myocardial infarction in young women: effect of female hormones on homocysteine levels. J Thromb Haemost 2(1):35–41
19. George J, Ahmed A, Patnaik M, Adler Y, Levy Y, Harats D, Gilburd B, Terrybery J, Shen GQ, Sagie A, Herz I, Snow P, Brandt J, Peter J, Shoenfeld Y (2000) The prediction of coronary atherosclerosis employing artificial neural networks. Clin Cardiol 23(6):453–456
20. She J, Deng Y, Wu Y, Xia Y, Li H, Liang X, Shi R, Yuan Z (2017) Hemoglobin A1c is associated with severity of coronary artery stenosis but not with long term clinical outcomes in diabetic and nondiabetic patients with acute myocardial infarction undergoing primary angioplasty. Cardiovasc Diabetol 16(1):97
21. Edlinger M, Wanitschek M, Dörler J, Ulmer H, Alber HF, Steyerberg EW (2017) External validation and extension of a diagnostic model for obstructive coronary artery disease: a cross-sectional predictive evaluation in 4888 patients of the Austrian Coronary Artery disease Risk Determination In Innsbruck by diaGnostic ANgiography (CARDIIGAN) cohort. BMJ Open 7(4):e014467
22. Vijaya Lakshmi SV, Naushad SM, Seshagiri Rao D, Kutala VK (2013) Oxidative stress is associated with genetic polymorphisms in one-carbon metabolism in coronary artery disease. Cell Biochem Biophys 67(2):353–361
23. Sun K, Song J, Liu K, Fang K, Wang L, Wang X, Li J, Tang X, Wu Y, Qin X, Wu T, Gao P, Chen D, Hu Y (2017) Associations between homocysteine metabolism related SNPs and carotid intima-media thickness: a Chinese sib pair study. J Thromb Thrombolysis 43(3):401–410
24. Nilsson BB, Lunde P, Grøgaard HK, Holm I (2018) Long-term results of high-intensity exercise-based cardiac rehabilitation in revascularized patients for symptomatic coronary artery disease. Am J Cardiol 121(1):21–26