STUDY ON THE RISK OF CORONARY HEART DISEASE IN MIDDLE-AGED AND YOUNG PEOPLE BASED ON MACHINE LEARNING METHODS: A RETROSPECTIVE COHORT STUDY

Mirvohidova Nazokat Mirbosit qizi

EMU University Student,Tashkent,Uzbekistan

Abstract: *Objective To identify coronary heart disease risk factors in young and middle-aged persons and develop a tailored risk prediction model.*

Keywords: Coronary heart disease, Young and middle-aged people, Logistic regression analysis, BP neural network, Random forest, XGBoost.

Methods

A retrospective cohort study was used in this research. From January 2024 to , 50 patients in the Department of Cardiology at a tertiary hospital in Anhui Province were chosen as research subjects. The research subjects were separated into two groups based on the results of coronary angiography performed during hospitalization (n = 201) and non-coronary heart disease (n = 352). R software (R 3.6.1) was used to analyze the clinical data of the two groups. A logistic regression prediction model and three machine learning models, including BP neural network, Extreme gradient boosting (XGBoost), and random forest, were built, and the best prediction model was chosen based on the relevant parameters of the different machine learning models.

Results

Univariate analysis identified a total of 24 indexes with statistically significant differences between coronary heart disease and non-coronary heart disease groups, which were incorporated in the logistic regression model and three machine learning models. The AUCs of the test set in the logistic regression prediction model, BP neural network model, random forest model, and XGBoost model were 0.829, 0.795, 0.928, and 0.940, respectively, and the F1 scores were

0.634, 0.606, 0.846, and 0.887, indicating that the XGBoost model's prediction value was the best.

Conclusion

The XGBoost model, which is based on coronary heart disease risk factors in young and middleaged people, has a high risk prediction efficiency for coronary heart disease in young and middle-aged people and can help clinical medical staff screen young and middle-aged people at high risk of coronary heart disease in clinical practice.

Introduction

Most of the previous epidemiological data came from the elderly (>65 years old), but due to obesity and poor lifestyle, the incidence rate of CHD increased rapidly in young and middle-aged patients. The Framingham Heart study reported the 10-year incidence rate of myocardial infarction (MI) in patients under 55 years old, 51.1/1,000 in men and 7.4/1,000 in women (Kannel & Abbott, 1984). (However, the literature on CHD and MI in young and middle-aged patients ≤65 years old is insufficient. The consequences of MI can be devastating, especially for young and middle-aged patients, because it has a greater potential impact on the patient's psychology, work ability and socio-economic burden. Previous studies have pointed out the differences between young and elderly MI patients. Compared with elderly MI patients, young MI patients have a larger proportion of men, a higher incidence of smoking and hyperlipidemia, a lower incidence of CHD, diabetes and hypertension, and their prognosis is better than that of elderly patients

Considering that there are many middle-aged and young people with CHD in recent years (Yunjun, Yanan & Quan, 2017), and as the main labor force of society, middle-aged and young people are at the core of work and family. If accompanied with CHD, it will have a great impact on their work and life, increase the economic burden and bring calm pressure to the society (Cuilu, 2022). Therefore, it is of great significance to screen out the middle-aged and young people with high risk of CHD and take active and effective prevention and control measures. In recent years, many

scholars have found that exploring new models of disease diagnosis based on machine learning algorithm has achieved good results in disease prediction and diagnosis (Dinh et al., 2019; Seo et al., 2019; Farran et al., 2019). Considering the harm of coronary heart disease in young and middle-aged people and the importance of early warning, this study used machine learning algorithm to establish an individual risk prediction model of coronary heart disease in young and middle-aged people, in order to provide an auxiliary diagnosis method for coronary heart disease in young and middle-aged people and reduce the risk of coronary heart disease in young and middle-aged people.

Data and methods Data sources

This study is a retrospective cohort study, 553 patients in the Department of Cardiology of a tertiary hospital in Anhui Province from January 2017 to January 2020 were taken as the research object, including 201 middle-aged and young people with coronary heart disease as the coronary heart disease group and 352 people without coronary heart disease as the non-coronary heart disease group. Diagnostic criteria of coronary heart disease: (1) symptoms of angina pectoris or MI attack; (2) ECG showed myocardial ischemia changes; (3) The operation items include coronary angiography. The coronary angiography shows that there is stenosis in at least one main branch of the left main artery, left anterior descending artery, left circumflex artery or right coronary artery, and the stenosis is more than 50%, and the patient is diagnosed as coronary heart disease after discharge. The medical ethics committee of the First Affiliated Hospital of the University of Science and Technology of China gave their approval to this study (ID: 2022-RE009). The subjects' informed consent was not required because this was a retrospective study and the data was analyzed anonymously.

Inclusion and exclusion criteria of the study population

Inclusion criteria: (1) The patient had no history of coronary heart disease; (2) Age 18-65 years old; (3) No mental illness. Exclusion criteria: (1) Combined with other acute and chronic infectious inflammation, cerebrovascular and renal vascular

diseases and tumors; (2) Persons with mental illness or unable to communicate normally; (3) Complicated with acute and chronic infectious inflammation, fracture, tumor, secondary hypertension or other serious physical diseases.

Index selection

The selected clinical data sources include patients' general data, cardiac ultrasound recording, laboratory examination results. General patient information includes complications (hypertension, diabetes, cerebral infarction), bad living habits (smoking, drinking), demographic data (education level, payment method of medical expenses, monthly family income, marital status, age, gender, body mass index (BMI), systolic blood pressure at admission, diastolic blood pressure at admission, mean arterial pressure at admission, pulse pressure at admission). Cardiac ultrasound recording includes left ventricular ejection fraction (LVEF), left ventricular enddiastolic dimension (LVEDD). Laboratory examination indicators includes thyroid-stimulating hormone (TSH), triiodothyronine (FT3), free thyroxine (FT4), very low density lipoprotein cholesterol (VLDL-C), low density lipoprotein cholesterol (LDL-C), high density lipoprotein cholesterol (HDL-C), triglyceride, cholesterol, blood calcium, blood sodium, blood potassium, blood carbon dioxide binding capacity, uric acid (UA), blood urea nitrogen (BUN), albumin (ALB), aspartate amino transferase (AST), alanine aminotransferase (ALT), platelet count (PLT), hemoglobin (HGB), red blood cell count (RBC), white blood cell count (WBC), Nterminal pro-brain natriuretic peptide (NT-proBNP), C-reactive protein (CRP), D-Dimer, Fasting blood glucose at admission. The above data are collected from the electronic medical record system of the First Affiliated Hospital of University of science and technology of China.

REFERENCES

1. Afifi M. Antecedant risk factors and their control in young patients with a first myocardial infarction. Singapore Medical Journal. 2006;47:554–556. - PubMed

2. Bae MH, Kim JH, Jang SY, Park SH, Lee JH, Yang DH, Park HS, Cho Y, Chae SC. Hyponatremia at discharge as a predictor of 12-month clinical outcomes in

hospital survivors after acute myocardial infarction. Heart and Vessels. 2017;32:126–133. doi: 10.1007/s00380016-0854-6. - DOI - PubMed

3. Bingrui W. Evaluation of serum albumin level in coronary artery disease in patients with ACS. 2018. Master's thesis, Suzhou University, Suzhou, China, 7–10.

4. Burkhardt K, Kirchberger I, Heier M, Zirngibl A, Kling E, von Scheidt W, Kuch B, Meisinger C. Hyponatraemia on admission to hospital is associated with increased long-term risk of mortality in survivors of myocardial infarction. European Journal of Preventive Cardiology. 2015;22:1419–1426. doi: 10.1177/2047487314557963. - DOI - PubMed

5. Che J, Li G, Shao Y, Niu H, Shi Y. An analysis of the risk factors for premature coronary artery disease in young and middle-age Chinese patients with hypertension. Experimental & Clinical Cardiology. 2013;18:89–92. - PMC - PubMed

6. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, Chen K, Mitchell R, Cano I, Zhou

T, Li M, Xie J, Lin M, Geng Y, Li Y, Yuan J. xgboost: extreme gradient boosting. The Comprehensive R Archive Network. 2021. https://CRAN.R-project.org/package=xgboost. [5 October 2021]. https://CRAN.R-project.org/package=xgboost R package version 1.4.1.1.

7. Chenghua Z, Shan H, Jingbo C, Li J, Luo W. Risk factors and characteristics of coronary artery disease in young and middle-aged patients with coronary heart disease. China Health Engineering. 2021;20(6):951–952+955. doi: 10.19937/j.issn.1671-4199.2021.06.023. - DOI

8. Chouhan L, Hajar HA, Pomposiello JC. Comparison of thrombolytic therapy for acute myocardial infarction in patients aged <35 and >55 years. The American Journal of Cardiology. 1993;71:157–159. doi: 10.1016/0002-9149(93)90731-q. - DOI - PubMed

9. Chunyan Y. Research on coronary heart disease screening model based on integrated feature selection[M.S.] Jinan: Shandong University; 2019. pp. 14–16.

10. Cuilu C. Investigation on health needs and influencing factors of young and middle-aged patients with coronary heart disease after percutaneous coronary intervention. Trace Elements and Health Research. 2022;39:1–4.