Journal of Kerman University of Medical Sciences

JKMU

Journal of Kerman University of Medical Sciences, 2021; 28 (2): 150-156

Rating and Comparison of Well-known Cardiovascular Risk Factors in Kermanian Male Patients Using Fuzzy Linear Regression

Milad Ahmadi Gohari, M.Sc.[1], Mansour Moazenzadeh, M.D.[2], Alireza Rashidinejad, M.D.[3], Abbas Bahrampour, Ph.D.[4,5]

1- Modeling in Health Research Center, Institute for Futures Studies in Health & Department of Biostatistics and Epidemiology, Faculty of Health, Kerman university of Medical Sciences, Kerman, Iran

2- Assistant Professor, Cardiovascular Research Center, Institute of Basic and Clinical Physiology Sciences, Kerman University of Medical Sciences, Kerman, Iran

3- Department of Cardiology, School of Medicine, Kerman University of Medical Sciences, Kerman, Iran.

4-Professor of Biostatistics, Modeling in Health Research Center, Institute for Futures Studies in Health & Department of Biostatistics and Epidemiology, Faculty of Health, Kerman university of Medical Sciences, Kerman, Iran

5-Adjunct Professor of Griffith University, Brisbane, QLD, Australia. (Corresponding author: Email: abahrampour@yahoo.com)

Received: 2 July, 2020 Accepted: 9 January, 2021

ARTICLE INFO

Article type:
Original Article

Keywords:
Fuzzy linear regression
Gensini score
Cardiovascular risk factors

Abstract

Background: Cardiovascular diseases are still among the most important causes of death in different countries. There are several risk factors for the onset of this disease. The rating of these risk factors is very important for informing the community and planning for the future.

Methods: Linear regression is one of the classic statistical methods that has many applications in medical sciences. When dealing with fuzzy data, it is not possible to use linear regression. The use of angiography to estimate the extent of congestion is associated with an estimate of more or less stenosis and the increase of atherosclerosis. Hence, this variable has been considered as a fuzzy variable. Fuzzy linear regression (FLR) was used to rank and compare the known risk factors for coronary artery occlusion.

Results: After analyzing the data by least squares FLR, the most important risk factors included Family history, history of diabetes, age, history of hypercholesterolemia, history of cigarette smoking, Body Mass Index and history of hypertension respectively.

Conclusion: When it is not possible to calculate the response variable or one of the independent variables examined in the model accurately, FLR, or to be more precise, regression in fuzzy environment can be a good alternative for conventional regression analysis.

Citation: Ahmadi Gohari M, Moazenzadeh M, Rashidinejad A, Bahrampour A. Rating and Comparison of Well-known Cardiovascular Risk Factors in Kermanian Male Patients Using Fuzzy Linear Regression. *Journal of Kerman University of Medical Sciences*, 2021; 28 (2): 150-156.

Introduction

Coronary artery disease (CAD) is one of the most important risk factors for mortality and morbidity in the World (1). According to W.H.O, "CVDs are the number one cause of death globally: more people die annually from CVDs than from any other causes. It was estimated that 17.9 million people died

from CVDs in 2016, representing 31% of all global deaths. Of these deaths, 85% were due to heart attack and stroke (2).

In Iran, CVD is also one of the most important causes of cardiovascular deaths (3). In general, studies have shown that in low and middle-income countries people are mostly affected by CVD during the working age compared with developed countries where a large statistics of CVD belongs to the retired individuals (4, 5). The incidence of mortality due to coronary heart disease in males at middle ages (30-64 years old) is 4 to 5 times that of women (6). So, cardiovascular disease imposes extra costs on the health system of countries and patients (7). Through preventing risk factors of major cardiac diseases and promoting healthy lifestyles in developing countries, their incidence can be reduced in the future.

Many factors are involved in creating cardiovascular diseases, some of which are inevitable like age, gender, heredity, but some of them can be controlled like: high blood pressure, elevated serum cholesterol, cigarette smoking, obesity, diabetes and etc. (8-10).

There are several ways to measure the extent of coronary artery stenosis. One of the most widely used is Gensini Score that is used to evaluate the severity of atherosclerosis (11, 12). Many factors such as the position and severity of arteries stenosis influence Gensini grading scale. In this quantitative scoring system, coronary stenosis is rated from zero to 4, and the involved parts are scored from 5 to 20. The multiplication of these two values is the patient's Gensini score of angiography (11).

In this scoring system, the patient who has the greater number of arteries stenosis and more dangerous position of stenosis achieves the higher score. This amount is identified by the specialist, but its measurement is not completely accurate.

Use of angiography to estimate CAD burden is challenged by under- or over-estimation of atherosclerotic narrowing and diffuseness, limitations in technique and disparities between lesion severity determined by angiography and true atherosclerosis burden (12). Therefore, there is usually no consensus on the Gensini scores among specialists. Therefore, this variable can be considered as an obscure or fuzzy variable.

Many of the events that are happening in the world are subject to some other events. One of the most important purposes of scientific researches especially medical researches is to detect the existence, absence, or the relationship between phenomena and also to predict based on the occurrence of various events. Regression analysis is one of the most important statistical methods, which is widely used for this purpose. However, regression analysis is subject to conditions. The most important of them is the exactness (not vagueness) of events. But, there are a lot of events which are not precisely measurable or we do not have access to the exact value of them. Therefore, they can be estimated such as the exact IQ, air pollution, the amount of water in a lake, factors affecting the sensitivity to a specific drug, the amount of coronary artery stenosis, the precise lifespan of an electronic chip, and etc. Hence, FLR, or to be more precise, regression in fuzzy environment can be a good alternative for conventional regression analysis. At first, fuzzy regression was examined by Tanaka et al. in 1982 and later it was expanded and considered by many researchers (13).

This research aimed to determine the relationship between the well-known risk factors of heart diseases and the severity of coronary artery stenosis (based on Gensini criteria) and to compare and rate the risk factors in male patients (between 35 to 70 years old) referred to Shafa Hospital in Kerman city and

underwent angiography for definite diagnosis of coronary lesion by using FLR.

Material and Methods

Data

Based on the rule of thumb, minimum of 10 events per the predictor variable is based for analysis (14). A group of 170 male patients aged between 35 to 70 years old referred to Shafa Hospital in Kerman city and underwent angiography for the diagnosis of coronary lesion was chosen as the sample size.

To collect data a checklist was used. The independent variables were: 1-Family history, 2- History of diabetes, 3-Age, 4-History of hypercholesterolemia, 5- History of smoking, 6-Body mass index, and 7- History of hypertension. The amount of coronary artery stenosis based on Gensini criterion, which is represented as approximate and fuzzy data, was considered as the response variable. The angiography video was seen by cardiologist and the Gensini score was calculated. To analyze the data, a program was written in software R.3.5.2 by researchers. In the present study, the amount of coronary artery stenosis (based on Gensini criterion) is assumed as symmetrical triangular fuzzy numbers that the ambiguity of each number is 10 percent of its exponent. Also, after standardizing the independent variables, we used fuzzy least squares regression. In this method, the coefficients of regression model are estimated based on definitions for the distance between fuzzy numbers.

Variables

Family history (FH): having the father or brother under the age of 55 and the mother or sister under the age of 65 who had been diagnosed with cardiovascular disease were considered as the family history (15).

History of cigarette smoking (CS): the duration of cigarette smoking (years or months) and the number of daily smoked cigarettes were asked.

History of hypertension disease (Hy): the length of time (years or months) that the patient has been diagnosed with hypertension and the related therapeutic interventions and medicines were determined.

History of diabetes disease (Di) and hypercholesterolemia disease (Ch): similar to the above variables have been determined.

Gensini score: the quantitative amount that shows the severity of coronary artery disease.

Body mass index (BMI): the body weight in kilograms (kg) divided by height in meters squared.

Fuzzy least squares regression method and deffuzification

In FLR, assuming that the response variable is fuzzy, there are different methods for fitting the model. We used the least squares method (16) for data analysis. For further details about pure topics, you can refer to the referenced articles. The basic idea of this approach, often referred to as possibilistic regression, is to minimize the fuzziness of the model by minimizing the total spread of the fuzzy coefficients (17).

In this method we used minimization of the Euclidean distance between the observed data and an adaptive model. After transforming the dependent variable and standardizing the independent variables, fuzzy regression equation was calculated.

Our general model is like below:

$$\tilde{Y} = \widetilde{A_0} + \widetilde{A_1}x_1 + \cdots \widetilde{A_n}x_n$$

Where $\tilde{y}$ is the fuzzy output, $\widetilde{A_i}$ is a fuzzy coefficients, and x_i is an n-dimensional non-fuzzy input vector. The fuzzy components were assumed to be triangular fuzzy numbers.

So, the coefficients of independent and dependent variables were calculated fuzzy. Therefore, in order to better interpret the coefficients, we should make the coefficients non-fuzzy. In recent years, various methods have been proposed for defuzzification. The most common is the gravity or centralized method. In this method, the weighted average of all variable values are obtained (weight of each output is its membership degree) (18).

Goodness of fits

To evaluate the goodness of fit, there is an indicator that can be used to calculate the goodness of fits. In this way, the adapting amount of every Y observation and the estimated value of the model are calculated. The more or less value of these numbers indicates the goodness or weak goodness of fit of the model respectively. According to the adaptive index, it is possible to identify the remote data (outliers). The data for which this index has a small value is considered as outliers (19).

Result

Descriptive results

Descriptive independent variables are shown in Table 1.

Table 1. Descriptive independent variables in the studied patients

Variables	Number (Percent)	Mean (S.D)
History of hypertension	42 (24.7)	1.2 (2.8)
History of cigarette smoking	84 (49.4)	11.3 (14.2)
History of diabetes	43 (25.3)	1.4 (3.5)
History of hypercholesterolemia	32 (18.8)	0.8 (2.3)
Family history	60 (35.5)	-
Age	-	54.7 (8.1)
BMI	-	24.6 (3.8)

About one quarter of the participants had a history of hypertension and diabetes, with an average disease duration of 1.2 and 1.4 years, respectively. Approximately half of the men participating in this study had a history of smoking with an average of about 11 smoked cigarettes per day. More than a third of the participants in the study had a history of heart disease. The average Body Mass of participants was 24.6.

Data analysis

We calculated FLR with transforming the dependent variable LN and standardizing the independent variables. In the end, after obtaining the parameters that are estimated as fuzzy, for better interpretation, it is possible to convert them into real numbers with defuzzy methods, so regression model is as follows:

Y=3.493+(0.06609*Age)+(0.0196*BMI)+(0.07129*FH)
+(0.00265*Hy)+(0.0227*CI)+(0.068*Di)+(0.0312*Ch)

As the coefficients show, the important risk factors are FH, Di, Age, Ch, CS, BMI and Hy, respectively.

The predictive power of the model

To assess the predictive power of the model, we compare the actual Gensini scores with estimated scores. Since both response variables were fuzzy values, we used the index to calculate the goodness of fit (19). Based on this index, it was found that for patients who did not have any risk factors, the goodness of fit was weak. The more is the history of risk factors, the more is goodness of fitting.

Discussion

The purpose of this study was to determine the effect and rating of each of the examined risk factors on the amount of coronary artery stenosis, and to estimate it based on fuzzy regression model and according to cultural, ethnic and diet characteristics of native patients. The results of our study indicated the importance of any of the examined risk factors of coronary artery disease based on Gensini criteria including family history, history of diabetes, age, history of hypercholesterolemia, history of cigarette smoking, body mass index and history of hypertension respectively. After fitting the model and examining goodness of fit, we concluded that FLR is an appropriate model to predict coronary artery stenosis based on the amount of having any of the cardiovascular diseases risk factors and considering Gensini score.

The risk factors studied in the present research and their effects on the cardiovascular diseases have been evaluated in many other studies. The family history of CVD was independently associated with premature CHD (odds ratio= 9.0) in a previous study (20). According to the American Heart Association (AHA): "adults with type 2 diabetes have a 2-to-4-fold higher risk for cardiovascular morbidity and mortality than adults without diabetes (21). Furthermore, the AHA deems diabetes is 1 of the 7 major controllable risk factors for cardiovascular disease (CVD)" (21). Another major Risk factor is age as incidence and prevalence of heart failure increase strikingly with age, with a lifetime risk of 20–46% (22, 23). A systematic review and meta-analysis was done in 2020and indicated that passive smoking was associated with a 12% higher increase in the risk of CVD mortality (24). In some studies, obesity has been considered as an independent risk factor for the early incidence of CAD (25, 26). According to the Framingham Heart Study, hypertension has a 2-fold increase in the development of heart failure in men (27).

Studies have been conducted in different countries to rank the risk factors. In these studies, a number of different risk factors were used compared to the present study. A study was conducted in 2014 on pooled data from 97 pooled prospective cohort studies that collectively enrolled 1.8 million participants between 1948 and 2005 (28). Their study compared the Hazard Ratios for the associations of physiological risk factors with Ischemic Heart Disease (IHD) and stroke between Western and Asian cohorts. The results showed that Hazard Ratios for IHD were Blood pressure, BMI, Serum cholesterol and Fasting Plasma Glucose (FPG) respectively for Asian cohorts and BMI, Blood pressure, Serum cholesterol and FPG, respectively for Western cohorts (29).

The reason for the difference between the results of our study and other studies can be attributed to the method of analysis, the study of different risk factors and the sample size and other factors such as race, socioeconomic status and diet.

Drawbacks

In the previous studies that have been conducted to identify risk factors of the disease, various results have been obtained based on different races, diets and different cultures. It means that, the results of the present study depend on racial characteristics, diets and even the culture of the studied subjects.

There are other well-known risk factors such as psychological factors (depression, stress) that can have a significant impact on heart disease, but it was not possible for us to include these variables.

Acknowledgement

This study was supported partially by Kerman University of Medical Sciences, Kerman, Iran; meanwhile, it is a part of Ahmadi Gohari's M.Sc. thesis. We also thank Ms. Hakimeh Safizadeh who helped to collect data and we appreciate the editor and reviewers for comments helping to improve the paper.

References

1. WHO. Cardiovascular diseases. [cited 2021 Feb 12] Available from: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1.
2. WHO. Cardiovascular diseases (CVDs). [cited 2021 Feb 12] Available from: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
3. Saadat S, Yousefifard M, Asady H, Moghadas Jafari A, Fayaz M, Hosseini M. The most important causes of death in Iranian population; a retrospective cohort study. Emerg (Tehran) 2015; 3(1):16-21.
4. Sadeghi M, Haghdoost AA, Bahrampour A, Dehghani M. Modeling the burden of cardiovascular diseases in Iran from 2005 to 2025: the impact of demographic changes. Iranian Journal of Public Health 2017; 46(4):506-16.
5. Sarrafzadegan N, Mohammmadifard N. Cardiovascular disease in Iran in the last 40 years: prevalence, mortality, morbidity, challenges and strategies for cardiovascular prevention. Arch Iran Med 2019; 22(4):204-10.
6. Bots SH, Peters SA, Woodward M. Sex differences in coronary heart disease and stroke mortality: a global assessment of the effect of ageing between 1980 and 2010. BMJ Glob Health 2017; 2(2):e000298.
7. Carter HE, Schofield D, Shrestha R. Productivity costs of cardiovascular disease mortality across disease types and socioeconomic groups. Open Heart 2019; 6(1):e000939.
8. Zdrenghea D, Guşetu G, Zdrenghea M, Cismaru G, Caloian B, Vaidean G, et al. CV RISK–A new relative cardiovascular risk score. Medical Hypotheses 2019; 132:109362.
9. Zhang Y, Cong H, Man C, Su Y, Sun H, Yang H, et al. Risk factors for cardiovascular disease from a population-based screening study in Tianjin, China: a cohort study of 36,215 residents. Ann Transl Med 2020; 8(7):444.
10. Psaltopoulou T, Hatzis G, Papageorgiou N, Androulakis E, Briasoulis A, Tousoulis D. Socioeconomic status and risk factors for cardiovascular disease: impact of dietary mediators. Hellenic J Cardiol 2017; 58(1):32-42.
11. Rampidis GP, Benetos G, Benz DC, Giannopoulos AA, Buechel RR. A guide for gensini score calculation. Atherosclerosis 2019; 287:181-3.
12. Neeland IJ ,Patel RS, Eshtehardi P, Dhawan S, McDaniel MC, Rab ST, et al. Coronary angiographic scoring systems: an evaluation of their equivalence and validity. Am Heart J 2012; 164(4):547-52.

13. Tanaka H. Fuzzy data analysis by possibilistic linear models. Fuzzy Sets and Systems 1987; 24(3):363-75.

14. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007; 165(6):710-8.

15. Scheuner MT, Whitworth WC, McGruder H, Yoon PW, Khoury MJ. Expanding the definition of a positive family history for early-onset coronary heart disease. Genet Med 2006; 8(8):491-501.

16. D'Urso P, Gastaldi T. A least-squares approach to fuzzy linear regression analysis. Computational Statistics & Data Analysis 2000; 34(4):427-40.

17. Shapiro AF. Fuzzy regression models. Article of Penn State University 2005; 102(2):373-83.

18. Ross TJ. Fuzzy Logic with Engineering Applications. 3th ed. UK: Wiley; 2010. p. 89-111.

19. Sadeghpour Gildeh B, Gine D. A goodness of fit index to reliability analysis in fuzzy model. [cited 2020 Oct 11] Available from: http://www.wseas.us/e-library/conferences/switzerland2002/papers/391.pdf.

20. Chacko M, Sarma PS, Harikrishnan S, Zachariah G, Jeemon P. Family history of cardiovascular disease and risk of premature coronary heart disease: a matched case-control study. Wellcome Open Res 2020; 5:70.

21. Hudspeth B. The burden of cardiovascular disease in patients with diabetes. Am J Manag Care 2018; 24(13 Suppl):S268-S72.

22. Gong FF, Coller JM, McGrady M, Boffa U, Shiel L, Liew D, et al. Age-related longitudinal change in cardiac structure and function in adults at increased cardiovascular risk. ESC Heart Fail 2020; 7(3):1344-61

23. Huffman MD, Berry JD, Ning H, Dyer AR, Garside DB, Cai X, et al. Lifetime risk for heart failure among white and black Americans: cardiovascular lifetime risk pooling project. J Am Coll Cardiol 2013; 61(14):1510-7.

24. Khoramdad M, Vahedian-azimi A, Karimi L, Rahimi-Bashar F, Amini H, Sahebkar AJIl. Association between passive smoking and cardiovascular disease: a systematic review and meta-analysis. IUBMB Life 2020; 72(4):677-86.

25. Formentini FS, Nagano FE, Neto FD, Adam EL, Fortes FS, da Silva LF. Coronary artery disease and body mass index: What is the relationship? Clinical Nutrition ESPEN 2019; 34:87-93.

26. Jazayeri MA, Waheed S, Shah Z, Parashara D, Gupta K. Impact of body mass index on the association of ankle-brachial index with all-cause and cardiovascular mortality: results from the national health and nutrition examination survey. Mayo Clin Proc Innov Qual Outcomes 2019; 3(4):409-17.

27. Tackling G, Borhade MB. Hypertensive heart disease. [cited 2021 Feb 28] Available from: https://www.ncbi.nlm.nih.gov/books/NBK539800/.

28. Global Burden of Metabolic Risk Factors for Chronic Diseases Collaboration (BMI Mediated Effects), Lu Y, Hajifathalian K, Ezzati M, Woodward M, Rimm EB, et al. Metabolic mediators of the effects of body-mass index, overweight, and obesity on coronary heart disease and stroke: a pooled analysis of 97 prospective cohorts with 1 8 million participants. Lancet 2014; 383(9921):970-83.

29. Tzoulaki I, Elliott P, Kontis V, Ezzati MJC. Worldwide exposures to cardiovascular risk factors and associated health effects: current knowledge and data gaps. Circulation 2016; 133(23):2314-33.