RESEARCH ARTICLE

Serum Biomarkers and Classification and Regression Trees Can Discriminate Symptomatic from Asymptomatic Carotid Artery Disease Patients

Aikaterini Trikouraki[1*] , Dido Yova[1], Abraham Pouliakis[2], Aris Spathis[2], Konstantinos G. Moulakakis[3] and Georgios Matsopoulos[1]

Abstract

Objective: To assess biomarkers between symptomatic and asymptomatic patients, and to construct a classification and regression tree (CART) algorithm for their discrimination.

Patients and Methods: 136 patients were enrolled. They were symptomatic (high risk) ($N = 82$, stenosis degree $\geq 50\%$, proven to be responsible for ischemic stroke the last six months) and asymptomatic (low risk) ($N = 54$, stenosis degree $\leq 50\%$). Levels of fibrinogen, matrix metalloproteinase-1 (MMP-1), tissue inhibitor of metalloproteinase-1 (TIMP-1), soluble intercellular adhesion molecule (SiCAM), soluble vascular cell adhesion molecule (SvCAM), adiponectin and insulin were measured on a Luminex 3D platform and their differences were evaluated; subsequently, a CART model was created and evaluated.

Results: All measured biomarkers, except adiponectin, had significantly higher levels in symptomatic patients. The constructed CART prognostic model had 97.6% discrimination accuracy on symptomatic patients and 79.6% on asymptomatic, while the overall accuracy was 90.4%. Moreover, the population was split into training and test sets for CART validation.

Conclusion: Significant differences were found in the biomarkers between symptomatic and asymptomatic patients. The CART model proved to be a simple decision-making algorithm linked with risk probabilities and provided evidence to identify and, therefore, treat patients being at high risk for cardiovascular disease.

Keywords: Multiplex technology, Serum biomarkers, Atherosclerosis, Statistical analysis, Cart

1 Introduction

Atherosclerotic vascular disease, which among others can lead to stroke, is a major cause of morbidity and mortality in Western countries [1]. This disease starts when the arteries endothelium becomes damaged, subsequently the accumulation of low density lipoproteins and the inflammation of the arterial wall becomes the first stage of atherosclerosis. This stage is similar for both the carotid and the coronary arteries [2]. Afterwards, plaque formation and, therefore, arterial stenosis and a high risk for embolism, follow.

Several serum biomarkers have been proposed to be associated with atherosclerotic plaque formation, progression, as well as with clinical outcome. Fibrinogen has

---

*Correspondence: katherinetrikouraki@gmail.com
[1] Biomedical Optics and Applied Biophysics Lab, Electro Optics and Electronic Materials, School of Electrical and Computer Engineering, National Technical University of Athens, Iroon Polytechniou 9, Zografou Campus, 15780 Athens, Greece
Full list of author information is available at the end of the article

Trikouraki *et al. Artery Res* (2021) 27:159–166

160

a role in the initial stages for the formation of atherosclerotic plaque [3] and prospective studies have confirmed a strong effect of raised fibrinogen in the progression of stroke and arterial disease [4]. Matrix metalloproteinase-1 (MMP-1) is expressed in human carotid artery disease and correlates with plaque instability [5]; the latter is largely responsible for atherosclerosis complications [6]. Instability is characteristic for plaques with high extracellular lipid content and an excess of macrophages in the cap [7], release of proteolytic enzymes (i.e., matrix metalloproteinase) by macrophages is suggested as a mechanism of cap erosion, since it can degrade various extracellular matrix components [8].

Tissue inhibitor of metalloproteinase-1 (TIMP-1), has been identified as involved in plaque formation [9]. TIMPs which are synthesized and secreted in atherosclerotic lesions, predominantly from macrophages and smooth muscle cells, contribute both to inflammatory state as well as the extracellular remodeling, which occur during several steps in atherogenesis and vascular remodeling [10]. Cellular adhesion molecules (CAMs), specifically, soluble intercellular adhesion molecule (SiCAM) and soluble vascular cell adhesion molecule (SvCAM), mediate the adhesion and leucocytes migration, steps that might have an important role in early atherogenesis [11, 12]. Focal expression of CAMs has been identified in atherosclerotic lesions [13]. All clinical and experimental studies relevant to adiponectin suggest that it is a critical vascular protective molecule and its reduction can contribute to vascular injury in disease associated with metabolic disorder [14]. Adiponectin is an adipose-specific plasma protein that has important roles in atherosclerosis [15]. Experimental studies have demonstrated that adiponectin has anti-inflammatory and anti-atherogenic properties. Moreover, high insulin levels are related to an increased cardiovascular risk, which is well known from several studies. A positive relationship between insulin levels and atherosclerotic events is reported [16].

Despite all the aforementioned biomarkers have been acknowledged to play an important role in atherosclerosis, their use in the clinical practice for decision-making is rather difficult. However, computer science and artificial intelligence enabled the development of computer-assisted systems for the support of clinical diagnosis or therapeutic and treatment decisions. Various classification techniques such as neural networks [17], discriminant analysis [18], classification and regression trees (CARTs) [19, 20] or genetic algorithms [21] have been used in medicine. Among the various decision support techniques, CARTs are an attractive machine learning approach to extract knowledge from data; and simultaneously construct easily understandable by physicians algorithms, being linked with probabilities.

In this study, there are two targets: (a) to assess the levels of fibrinogen, matrix metalloproteinase-1 (MMP-1), tissue inhibitor of metalloproteinase (TIMP-1), soluble intercellular adhesion molecule (SiCAM), soluble vascular cell adhesion molecule (SvCAM), adiponectin and insulin between symptomatic vs. asymptomatic patients and (b) to construct and evaluate a CART model for the discrimination of these two groups of patients as a potential algorithm to be used in clinical practice.

2 Patients and Methods

2.1 Patient's Cohort

In the present investigation, randomly selected patients with established carotid atherosclerosis and subjects without apparent atherosclerotic manifestations were enrolled after informed consent. 136 patients, from the Vascular Surgery Clinic or the Cardiology Clinic of the University General Hospital "Attikon", were enrolled in the study. The study protocol conformed to the ethical guidelines of the 1975 Helsinki Declaration and was approved by the Ethics Committee of "Attikon" University Hospital. Patients were subsequently examined and assigned into the following groups [22]:

1. Symptomatic (high risk) group (stenosis degree $\geq 50\%$, which is proven to be responsible for ischemic stroke in the last six months): the number of patients in this group was 82.
2. Asymptomatic (low risk) group (stenosis degree $\leq 50\%$): it consisted of 54 individuals to identify differences from the patients with symptoms.

Moreover, it is well known that atherosclerosis can lead to stroke, also, this disease starts when the arteries endothelium becomes damaged; therefore, arterial stenosis, in a relatively high percentage, happens [1, 2].

2.2 Clinical Data Collection and Imaging Tests

Patients demographic data were recorded: including gender, age, smoking habits, body mass index (BMI, underweight < 18.5; normal weight 18.5–25; overweight 25–30; obese > 30), systolic blood pressure (Systolic BP, normal 100–140 mmHg), triglycerides (normal < 150 mg/dL; mildly high 150–199 mg/dL; high 200–499 mg/dL; very high > 500 mg/dL) white blood cells (WBC, average normal range 4500–10,000 counts/mm^3). Blood pressure was measured twice with the patient seated for at least 15 min and with an intermediate time interval of 5 min between the measurements. The average of the measurements was evaluated and recorded, without any differences between the right and the left arm.

To define the stenosis degree, all participants underwent carotid ultrasound followed by image analysis at

the study initiation. The ultra-sonographic test was performed by one experienced vascular surgeon to avoid inter-observer variability, using a linear transducer signal of 12 MHz (General Electric LogiqE, Riverside, USA). The patients were in a supine position with a slight stretch of the head, to be able to display the ipsilateral carotid in an appropriate suitable longitudinal and lateral projection. Both carotids, were tested with an angle between head and transducer < 60° and the patient being in apnea. To ensure the reproducibility of carotid measurements, a carotid ultrasound assessment protocol was developed, by the University of Athens team, based on co-evaluation of sonographic and angiographic findings. In that way, the longitudinal ultrasound imaging was taking place in standard view of the carotid, while all other parameters were stable (imaging modality, B-mode; dynamic range, 60 dB; persistence, low; frame rate, higher than 25 frames/s). To have comparability of the measured parameters between the groups, anatomical areas of interest were predetermined based on guide points that included 3 cm length of the proximal internal carotid artery, carotid bulb and 1 cm length of the common carotid artery. Atherosclerotic lesions were distributed in the 3 carotid regions as follows: (a) carotid division (50%), (b) internal carotid artery (30%), and (c) common carotid artery (20%). In addition, to avoid inter-observer variability, all patients were evaluated by a single and experienced vascular surgeon.

2.3 Biomarkers

For all patients, blood sampling was performed in the morning (8:00 and 10:00 a.m.) after an overnight fast. Serum and plasma were isolated after centrifugation at 700 g. Samples were stored at − 80 °C. Levels of all proteins, except fibrinogen, were measured in serum using multiplex analysis. Using commercially available kits, inflammation related biomarkers were selected. This included measurement of plasma concentrations of fibrinogen, matrix metalloproteinase-1 (MMP-1), tissue inhibitor of metalloproteinase (TIMP-1), soluble intercellular adhesion molecule (SiCAM), soluble vascular cell adhesion molecule (SvCAM), adiponectin and insulin (EMD Millipore Corporation, Darmstadt, Germany). Bead assays were performed according to the manufacturer's protocol (Millipore, Billerica, MA, USA) and were analyzed on a Luminex 3D platform (Luminex Corp, Austin, TX, USA).

Intra-assay precision for all proteins was generated from the mean of the %CV's from eight reportable results across two different concentrations of analytes in a single assay; thus for an overnight protocol, the values were: < 15%CV for fibrinogen, ≤ 10%CV for TIMP-1, < 15%CV for SiCAM and SvCAM, 2%CV for adiponectin and < 10% for insulin; while for a 2-h protocol, the values were: 2.6%CV for MMP-1. Inter-assay precision was generated from the mean of the %CVs across two different concentrations of analytes and across 4 different assays. The values were: < 20% CV for fibrinogen, ≤ 10%CV for TIMP-1, < 20%CV for SiCAM and SvCAM, 10%CV for adiponectin, < 15% for insulin and 8.4%CV for MMP-1, respectively.

2.4 Statistical Analysis

The variation of the biomarkers' levels between symptomatic and asymptomatic patients was performed by the independent-samples *t* test. Results were reported as mean value ± SD. The distributions were tested for normality by Kolmogorov–Smirnov analysis. A *p* value < 0.05 was considered to be statistically significant. SPSS statistics software (version 22.0; SPSS Inc., Chicago, IL, USA) was used both for statistical analysis and CART construction.

2.5 Classification and Regression Tree (CART)

Classification and regression tree [19] is a recursive, partitioning, machine learning technique that builds tree-like structures for predicting or discriminating continuous variables (hence regression) or categorical variables (classification). In this study, the produced CART tree-like algorithm, created a set of if–then logical/split rules, eventually allowing assignment of patients as symptomatic or asymptomatic. The probability of a case belonging to one of the two categories was also provided by the CART. The CART was composed of nodes (i.e., points where decisions are made), moreover nodes had a hierarchy of layers: the first layer had only one node ("root" node), nodes in subsequent layers were linked with nodes in two other layers (parent and children nodes, called "branches"), while this hierarchical structure ends with terminal nodes (having only parents but not children, called "leafs").

The CART was built in a recursive manner on the basis of the supplied clinical characteristics and serum biomarkers. During each recursion, statistical measures for all supplied parameters were calculated, subsequently the parameter that allows better separation of the cases was identified and used to separate the case as symptomatic or asymptomatic (i.e., a new if–then rule was created), moreover two (or more) new children nodes were created. The patients were distributed to each one of the "children" nodes, which contained the number of patients being symptomatic and asymptomatic as well as the relevant probability. Subsequently, and if each child node contained both symptomatic and asymptomatic cases, the algorithm was repeated for each new child node that was produced. Once again the parameter allowed better

Trikouraki *et al. Artery Res* (2021) 27:159–166

162

separation of cases, was statistically identified and was used to create a new if–then rule and new "children" nodes.

The CART minimum parent size was set to 10, the minimum child size was set to five and the maximum allowed depth was set to eight, the QUEST growing algorithm was employed, while a tenfold cross-validation was performed. Note that these limits were not reached, the selected values are typical values and rather heuristic.

3 Results

The mean age was higher than 70 years for both symptomatic and asymptomatic patients, specifically the age for the symptomatic patients was 72.3 ± 9.7 and for the asymptomatic group 70.4 ± 9.1 without a statistically significant difference ($p = 0.264$) and indicative that the two populations were matched in their age, similarly 57/82 (69.5%) of the symptomatic patients were men and 31/54 (57.4%) of the asymptomatic ($p = 0.148$); therefore, the population composition had no difference in relation to the gender. In addition, smoking has been evaluated as

a risk factor with results of 61.0% of the symptomatic group and 54.8% of the asymptomatic one, without a statistical significant difference ($p = 0.373$). Symptomatic patients experienced higher BMI, systolic blood pressure (BP), triglycerides and white blood cell counts (WBC) (in all cases $p < 0.05$, see Table 1). It is obvious that BMI and systolic blood pressure between the two groups, present a significant difference, as p value indicates, making them discriminant factors between symptomatic and asymptomatic patients.

Serum biomarkers' values from patients of symptomatic and asymptomatic group appear in Table 2. Levels of fibrinogen, MMP-1, TIMP-1, SiCAM, SvCAM and insulin were significantly increased in symptomatic patients compared to asymptomatic, while levels of adiponectin presented an increase in asymptomatic patients ($p < 0.05$).

The constructed CART model shows the results along with the category where every node "belongs", the number of correctly assigned cases in each category, as well as, the relevant percentages (within each individual node, see Fig. 1). From the seven proteins and the clinical characteristics provided as inputs to the CART, the most important (as these were identified during training) were: MMP-1, fibrinogen, insulin and TIMP-1 (Fig. 1), since these are the only ones remained in the CART.

As an example of the CART logic, the most important paths with a greater number of incidents were (Fig. 1): (a) the path with the proteins MMP-1, MMP-1, insulin and (b) the path with the proteins MMP-1, fibrinogen, insulin. According to the first path: if the levels of MMP-1 (the first parameter examined) were > 4375 then it was required another examination (MMP-1), in this situation if the levels of MMP-1 were > 5404, then the levels of insulin were controlled and if they were > 236, the case could be characterized as symptomatic with probability 96.7%. On the other hand, if the levels of MMP-1 were ≤ 4375, then it was required another step (fibrinogen) and if the

Table 1 Clinical characteristics of symptomatic vs. asymptomatic patients

Variables	Clinical characteristics		p value
	Symptomatic patients (number of patients: $n = 82$)	Asymptomatic patients (number of patients: $n = 54$)	
Body Mass Index (BMI) (kg/m^2)	26.9 ± 3.8	24.2 ± 2.3	**< 0.001**
Systolic Blood Pressure (BP) (mmHg)	141.9 ± 22.4	131.7 ± 13.2	**0.007**
Triglycerides (mg/dl)	227.0 ± 85.6	85.7 ± 47.1	**< 0.001**
White Blood Cells (WBC) (thousands/mm^3)	7.1 ± 1.8	5.4 ± 1.2	**< 0.001**

Data are: means ± SD

Table 2 Serum biomarkers' values of symptomatic vs. asymptomatic patients

Variables	Patient group		p value
	Symptomatic (number of patients: $n = 82$)	Asymptomatic (number of patients: $n = 54$)	
Fibrinogen (ng/ml)	$651,390 \pm 215,152$	$577,181 \pm 208,717$	**< 0.001**
Matrix MetalloProteinase (MMP-1) (pg/ml)	5503 ± 2271	4687 ± 2179	**< 0.001**
Tissue Inhibitor of MetalloProteinase (TIMP-1) (pg/ml)	$153,949 \pm 26,480$	$146,124 \pm 23,101$	**< 0.001**
Soluble interCellular Adhesion Molecule (SiCAM) (ng/ml)	3804 ± 754	1403 ± 332	**< 0.05**
Soluble vascular Cell Adhesion Molecule (SvCAM) (ng/ml)	5983 ± 2201	3273 ± 1107	**< 0.05**
Adiponectin (pg/ml)	$72,485 \pm 12,507$	$87,672 \pm 16,439$	**< 0.05**
Insulin (pg/ml)	295 ± 107	270 ± 89	**< 0.001**

Data are: means ± SD

Trikouraki *et al. Artery Res* (2021) 27:159–166

163

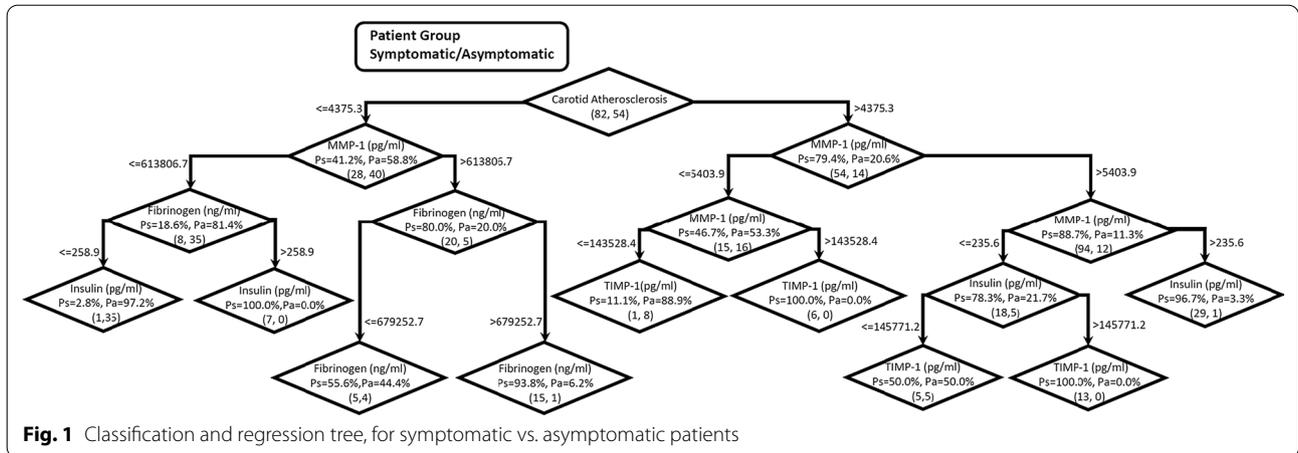**Fig. 1** Classification and regression tree, for symptomatic vs. asymptomatic patients

Table 3 Classification of observed versus assigned by the cart values, for symptomatic and asymptomatic patients

Real category (observed)	Classification		
	Assigned category (by the CART)		
	Symptomatic	Asymptomatic	Percent correct (%)
Symptomatic	80	2	97.6
Asymptomatic	11	43	79.6
Overall %	66.9%	33.1%	90.4

Table 4 Classification for real and assigned category in training and test set

Sample	Classification			
	Real category (observed)	Assigned category (by the CART)		
		High risk	Low risk	Percent correct (%)
Training	High risk	50	4	92.59
	Low risk	7	29	80.56
	Overall percentage	63.33%	36.67%	87.78
Test	High risk	25	3	89.29
	Low risk	4	14	77.78
	Overall percentage	63.04%	36.96%	84.78

levels of fibrinogen were ≤ 613,807, insulin was controlled, if the levels of insulin were ≤ 259, the case was characterized as asymptomatic with probability 97.2%. In a similar manner, the CART can be used to categorize a patient, symptomatic vs asymptomatic, from the important biomarkers along with a probability for all cases.

Cumulatively, the correlation matrix (Table 3) of the CART, representative of the performance, presents the percentage among the real category (observed) and the assigned category. The performance of this approach was: specificity: 79.6% (95% CI 67.1–88.2%), sensitivity: 97.6% (95% CI 91.5–99.3%). In the symptomatic group, 80 patients were accurately assigned as symptomatic (true-positive value), and 2 were falsely assigned as asymptomatic (false-positive value). On the other hand, in asymptomatic group the true negative value was 43 and the false negative value was 11. In addition to the tenfold cross-validation, the population was split into training and test sets, specifically about 2/3 of the data were randomly selected (respecting the distribution of cases into high and low risk) to train a CART and the remaining 1/3 (test set) were used for CART

validation. The results are presented in Table 4, enhancing the results with tenfold cross-validation technique. In the training set, the performance of this approach was: specificity: 80.6% (95% CI 65.0–90.3%), sensitivity: 92.59% (95% CI 82.4–97.1%) and in the test set: specificity: 77.8% (95% CI 54.8–91.0%), sensitivity: 89.3% (95% CI 72.8–96.3%).

While implementing the training and test method in CART, we split the population of all patients (i.e. individuals with stenosis degree ≥ 50%), into two main categories including approximately the 60% and 40% of the population respectively; the former class of patients contained the ones responsible for ischemic stroke and the latter those to whom no incidence occurred. The overall pool is the original 136 patients and were added 8 more patients who had stenosis degree ≥ 50% but did not present stroke events and the results are presented in Table 5. The performance of this model in the training set was: specificity: 83.3% (96% CI 66.4–92.7%),

Trikouraki *et al. Artery Res* (2021) 27:159–166

164

Table 5 Classification for real and assigned category in training and test set

Sample	Classification			
	Real category (observed)	Assigned category (by the CART)		
		Stroke	No stroke	Percent correct (%)
Training	Stroke	43	4	91.5
	No stroke	5	25	83.3
	Overall percentage	62.3%	37.7%	88.3
Test	Stroke	27	8	77.1
	No stroke	6	26	81.3
	Overall percentage	49.3%	50.7%	79.1

sensitivity: 91.5% (96% CI 81.1–96.6%) and in the test set: specificity: 81.3% (96% CI 64.7–91.1%), sensitivity: 77.1% (96% CI 65.6–91.4%).

4 Discussion

The results of this study showed significant differences in protein levels, between patients with and without symptoms ($p < 0.05$). The levels of the measured biomarkers in symptomatic patients were higher from those in asymptomatic patients, except adiponectin. More specifically, levels of fibrinogen, matrix metalloproteinase-1 (MMP-1), tissue inhibitor of metalloproteinase (TIMP-1), soluble intercellular adhesion molecule (SiCAM), soluble vascular cell adhesion molecule (SvCAM), and insulin were significantly increased in symptomatic patients compared to asymptomatic ($p < 0.05$).

Elevated levels of various biomarkers of inflammation have been linked to an increased risk of ischemic cardiovascular events. Our results are consistent with that documented by Zureik et al. [9] who presented a potential involvement of TIMP-1 in plaque formation (i.e., higher levels of TIMP-1 in patients with carotid atherosclerosis). Nikkari [5] documented that MMP-1 is expressed in human carotid artery disease and there is a correlation among this protein and plaque instability, also consistent with this research. Fibrinogen has also been proposed, as a marker of atheroma, and its high concentrations have been associated with ischemic heart disease and subclinical carotid artery disease [3]. Ginsberg [23] investigated insulin and the association with carotid artery disease; a significant increase of the risk for symptomatic carotid artery disease was also verified by our results. Ducimietere et al. [24], Pyorala [25] and Orchand et al. [16] reported a positive relationship between insulin

levels and atherosclerotic events, also compatible with our results.

There are two popular methods to evaluate the robustness of classifiers: (a) splitting the dataset into a part that is used for training and using another part of the data used for testing, this is useful when there is a very large dataset; therefore, the test dataset is able to provide a good estimation of system performance [19, 26] and (b) the k-fold method, i.e., splitting the dataset into k-subsets, each subset is held and the model is trained using the remaining subsets, this is repeated and an average accuracy estimate is produced. The second approach is considered a robust method for estimating accuracy, and specifically the amount of bias in the estimate has one of the popular values, number 10 [27–29]. We have selected the tenfold cross-validation due to the rather limited number of cases in our study.

Data through cross-validation from the CART demonstrated interesting results for the identification of both symptomatic and asymptomatic patients. In case that a CART "leaf" contains both symptomatic and asymptomatic cases, then the case is characterized according to the group that has the higher probability. The obtained results showed that the CART discrimination was accurate for about 97% of symptomatic patients and approximately 80% of asymptomatic. If the levels of MMP-1 were > 5404 and if the levels of insulin were > 236, then a patient can be assigned as symptomatic with confidence level 96.7%.

The proposed methodology has an important advantage. Notably despite the CART was supplied with 10 features (seven protein and three clinical), the final architecture (Fig. 1) was stabilized to provide results with just four biomarkers (MMP-1, fibrinogen, insulin and TIMP-1); therefore, not all protein measurements, neither all clinical tests are necessary to be performed; this introduces reduction of expenses in consumables and human effort. From the CART (Fig. 1), there are two paths that the greater number of incidents are concentrated, first the path with the proteins MMP-1, MMP-1, insulin, and second, the path with the proteins MMP-1, fibrinogen, insulin. As an example, if the levels of MMP-1 (the first parameter examined) were > 4375, then it was required another examination (MMP-1), in this situation if the levels of MMP-1 were > 5404, then the levels of insulin were controlled and if they were > 236, the case could be characterized as symptomatic with confidence level 96.7%. On the other hand, if the levels of MMP-1 were ≤ 4375, then it was required another step (fibrinogen) and if the levels of fibrinogen were ≤ 613,807, then insulin was controlled, if the levels of insulin were ≤ 259 then the case could be assigned as asymptomatic with confidence level 97.2%.

CARTs represent a very popular technique being used in the past by this research team to solve numerous medical classification problems [30–32]. Their main advantages include:

a) they produce simple to understand and interpret algorithms, that can be used even without the use of a computer

b) they are non-parametric, thus there is no need to check if data meet special conditions required for statistical tests, such as normality

c) they can handle both numeric and categorical data simultaneously

d) the produced algorithms might "mirror" and divert human decision processes

e) the produced rules in the tree-like structures contain important knowledge and can be used for the discovery of new "phenomena" and extraction of new knowledge

f) they are efficient when large data sets are available and required, thus there are no required specialized computing resources

g) it is possible to validate the produced models via statistical tests or simply by counting the number of successfully classified cases and the number of "missed" cases.

This is the first study attempting to use this CART tool in carotid artery disease, to investigate and discriminate symptomatic from asymptomatic patients. The objective of this paper was to assess the significant differences in protein levels between patients with and without symptoms, as well as, to evaluate whether there could be an accurate discrimination between these two cohorts: symptomatic vs asymptomatic. Among all protein levels, there was a statistically significant difference between symptomatic and asymptomatic group. Additionally, CARTs demonstrated results of high discrimination accuracy among symptomatic and asymptomatic patients and identified the most important characteristics from the inputs. So, identifying and studying specific circulating adipokines in combination with demographic data was important.

This study has limitations. The number of patients is relatively low; moreover, in the literature there are reported various other biomarkers being associated with carotid artery disease, these were not assessed in this study, despite they could have the potential to act as discriminant factors. In contrast, our study used a relatively large number of well-known carotid plaque inflammatory biomarkers and gave emphasis to the role of a CART model that can act in a combinatory method and separate symptomatic or asymptomatic patients. Another

study limitation is relevant to the patient ages. It is well known that the incidence of stroke and thus symptomatic carotid artery stenosis is higher in men until advanced age, and actually there is a higher incidence in women after 85 years old [33]. Since our population has a mean age around 70, there were more men than women with symptomatic stenosis; moreover, it is expected a higher percentage of men in the asymptomatic group [34]. In this study, the gender was not correlated with the symptomatic or asymptomatic patients. Finally, in this study, we have not examined the influence of administered drugs in the studied parameters due to the relatively small sample size.

5 Conclusion

Concluding, the application of a CART prognostic model for carotid artery disease provided a simple decision algorithm linked with probabilities. The prognostic value in symptomatic patients was high > 97%, and this tool can provide encouraging evidence to identify and treat symptomatic from asymptomatic patients. This discrimination could be important in the clinical practice, since non-symptomatic patients may already have experienced unnoticed symptoms that can be identified using imaging techniques, thus, a prospective trial in the clinical environment could evaluate the CART value in practice. Such trial requires development of software, that would hide CART details, with the help of a user-friendly computer interface.

Declarations

Conflict of interest
The authors report no conflicts of interest.

Ethical approval
The study protocol conformed to the ethical guidelines of the 1975 Helsinki Declaration and was approved by the Ethics Committee of "Attikon" University Hospital.

Consent for publication
All authors have read and approved the manuscript.

Trikouraki *et al. Artery Res* (2021) 27:159–166

166

Data Availability

Author details

[1]Biomedical Optics and Applied Biophysics Lab, Electro Optics and Electronic Materials, School of Electrical and Computer Engineering, National Technical University of Athens, Iroon Polytechniou 9, Zografou Campus, 15780 Athens, Greece. [2]Second Department of Pathology National and Kapodistrian University of Athens, "Attikon" University Hospital, Rimini 1, 12462 Athens, Greece. [3]Department of Vascular Surgery, National and Kapodistrian University of Athens, "Attikon" University Hospital, Rimini 1, 12462 Athens, Greece.

References

1. Huff MW, Daugherty A, Lu H. Atherosclerosis. In: Ridgway N, McLeod R, editors. Biochemistry of lipids, lipoproteins and membranes. 6th ed. Amsterdam: Elsevier; 2016. p. 519–48.
2. Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. Nature. 2011;473:317–25.
3. Meade TW. Fibrinogen and cardiovascular disease. J Clin Pathol. 1997;50:13–5.
4. Ernst E, Resch KL. Fibrinogen as a cardiovascular risk factor: a meta-analysis and review of the literature. Ann Intern Med. 1993;118:956–63.
5. Nikkari ST, O'Brien KD, Ferguson M, Hatsukami T, Welgus HG, Alpers CE, et al. Interstitial collagenase (MMP-1) expression in human carotid atherosclerosis. Circulation. 1995;92:1393–8.
6. Davies MJ, Thomas AC. Plaque fissuring–the cause of acute myocardial infarction, sudden ischaemic death, and crescendo angina. Br Heart J. 1985;53:363–73.
7. Davies MJ, Richardson PD, Woolf N, Katz DR, Mann J. Risk of thrombosis in human atherosclerotic plaques: role of extracellular lipid, macrophage, and smooth muscle cell content. Br Heart J. 1993;69:377–81.
8. Richardson PD, Davies MJ, Born GV. Influence of plaque configuration and stress distribution on fissuring of coronary atherosclerotic plaques. Lancet. 1989;2:941–4.
9. Zureik M, Beaudeux JL, Courbon D, Benetos A, Ducimetiere P. Serum tissue inhibitors of metalloproteinases 1 (TIMP-1) and carotid atherosclerosis and aortic arterial stiffness. J Hypertens. 2005;23:2263–8.
10. Galis ZS, Khatri JJ. Matrix metalloproteinases in vascular remodeling and atherogenesis: the good, the bad, and the ugly. Circ Res. 2002;90:251–62.
11. Ross R. The pathogenesis of atherosclerosis: a perspective for the 1990s. Nature. 1993;362:801–9.
12. Fotis L, Agrogiannis G, Vlachos IS, Pantopoulou A, Margoni A, Kostaki M, et al. Intercellular adhesion molecule (ICAM)-1 and vascular cell adhesion molecule (VCAM)-1 at the early stages of atherosclerosis in a rat model. In Vivo. 2012;26:243–50.
13. Davies MJ, Gordon JL, Gearing AJ, Pigott R, Woolf N, Katz D, et al. The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. J Pathol. 1993;171:223–9.
14. Ouchi N, Ohishi M, Kihara S, Funahashi T, Nakamura T, Nagaretani H, et al. Association of hypoadiponectinemia with impaired vasoreactivity. Hypertension. 2003;42:231–4.
15. Ouchi N, Kihara S, Arita Y, Maeda K, Kuriyama H, Okamoto Y, et al. Novel modulator for endothelial adhesion molecules: adipocyte-derived plasma protein adiponectin. Circulation. 1999;100:2473–6.
16. Orchard TJ, Eichner J, Kuller LH, Becker DJ, McCallum LM, Grandits GA. Insulin as a predictor of coronary heart disease: interaction with apolipoprotein E phenotype. A report from the multiple risk factor intervention trial. Ann Epidemiol. 1994;4:40–5.
17. Pouliakis A, Karakitsou E, Margari N, Bountris P, Haritou M, Panayiotides J, et al. Artificial Neural networks as decision support tools in cytopathology: past, present, and future. Biomed Eng Comput Biol. 2016;7:1–18.
18. Datta S. Classification of breast cancer versus normal samples from mass spectrometry profiles using linear discriminant analysis of important features selected by random forest. Stat Appl Genet Mol Biol. 2008;7: Article7.
19. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. Belmont: Wadsworth International Group; 1984.
20. Karakitsos P, Pouliakis A, Meristoudis C, Margari N, Kassanos D, Kyrgiou M, et al. A preliminary study of the potential of tree classifiers in triage of high-grade squamous intraepithelial lesions. Anal Quant Cytol Histol. 2011;33:132–40.
21. Ghaheri A, Shoar S, Naderan M, Hoseini SS. The applications of genetic algorithms in medicine. Oman Med J. 2015;30:406–16.
22. Kadoglou NP, Lambadiari V, Gastounioti A, Gkekas C, Giannakopoulos TG, Koulia K, et al. The relationship of novel adipokines, RBP4 and omentin-1, with carotid atherosclerosis severity and vulnerability. Atherosclerosis. 2014;235:606–12.
23. Ginsberg HN. Insulin resistance and cardiovascular disease. J Clin Investig. 2000;106:453–8.
24. Ducimetiere P, Eschwege E, Papoz L, Richard JL, Claude JR, Rosselin G. Relationship of plasma insulin levels to the incidence of myocardial infarction and coronary heart disease mortality in a middle-aged population. Diabetologia. 1980;19:205–10.
25. Pyorala K. Relationship of glucose tolerance and plasma insulin to the incidence of coronary heart disease: results from two population studies in Finland. Diabetes Care. 1979;2:131–41.
26. Mena E, Bolte G, Advance Gender Study G. CART-analysis embedded in social theory: a case study comparing quantitative data analysis strategies for intersectionality-based public health monitoring within and beyond the binaries. SSM Population Health. 2021;13:100722.
27. Ghiasi MM, Zendehboudi S, Mohsenipour AA. Decision tree-based diagnosis of coronary artery disease: CART model. Comput Methods Prog Biomed. 2020;192:105400.
28. Shim EJ, Yoon MA, Yoo HJ, Chee CG, Lee MH, Lee SH, et al. An MRI-based decision tree to distinguish lipomas and lipoma variants from well-differentiated liposarcoma of the extremity and superficial trunk: classification and regression tree (CART) analysis. Eur J Radiol. 2020;127:109012.
29. Brownlee J. Machine learning mastery with R,Packt, O'Reilly 2016.
30. Pergialiotis V, Pouliakis A, Parthenis C, Damaskou V, Chrelias C, Papantoniou N, et al. The utility of artificial neural networks and classification and regression trees for the prediction of endometrial cancer in postmenopausal women. Public Health. 2018;164:1–6.
31. Pouliakis A, Karakitsou E, Chrelias C, Pappas A, Panayiotides I, Valasoulis G, et al. The application of classification and regression trees for the triage of women for referral to colposcopy and the estimation of risk for cervical intraepithelial neoplasia: a study based on 1625 cases with incomplete data from molecular tests. Biomed Res Int. 2015;2015:914740.
32. Pouliakis A, Margari C, Margari N, Chrelias C, Zygouris D, Meristoudis C, et al. Using classification and regression trees, liquid-based cytology and nuclear morphometry for the discrimination of endometrial lesions. Diagn Cytopathol. 2014;42:582–91.
33. Turtzo LC, McCullough LD. Sex differences in stroke. Cerebrovasc Dis. 2008;26:462–74.
34. de Weerd M, Greving JP, de Jong AW, Buskens E, Bots ML. Prevalence of asymptomatic carotid artery stenosis according to age and sex: systematic review and metaregression analysis. Stroke J Cerebral Circ. 2009;40:1105–13.