A Machine Learning Model for the prediction of the progression of carotid arterial stenoses

Panagiotis K. Siogkas, Dimitrios S. Pleouras, Vasilis D. Tsakanikas, Vassiliki T. Potsika, Antonis Sakellarios, Evdokia Karamouzi, Foteini Lagiou, George Charalampopoulos, George Galyfos, Fragiska Sigala, Igor Koncar and Dimitrios I. Fotiadis, *Fellow, IEEE*

***Abstract*— Atherosclerotic carotid plaque development results in a steady narrowing of the artery lumen, which may eventually trigger catastrophic plaque rupture leading to thromboembolism and stroke. The primary cause of ischemic stroke in the EU is carotid artery disease, which increases the demand for tools for risk stratification and patient management in carotid artery disease. Additionally, advancements in cardiovascular modeling over the past few years have made it possible to build exact three-dimensional models of patient-specific primary carotid arteries. Computational models then incorporate the aforementioned 3D models to estimate either the development of atherosclerotic plaque or a number of flow-related parameters that are linked to risk assessment. Computational models make it feasible to accurately describe the blood flow and the basic atherosclerotic process within the artery wall, enabling the prediction of future lumen stenoses, plaque areas, and risk prediction. This work presents an attempt to provide a carotid artery stenosis prognostic model, utilizing non-imaging and imaging data, as well as simulated hemodynamic data. The overall methodology was trained and tested on a dataset of 41 cases with 23 carotid arteries with stable stenosis and 18 carotids with increasing stenosis degree. The highest accuracy of 71% was achieved using a neural network classifier. The novel aspect of our work is the definition of the problem that is solved, as well as the amount of simulated data that are used as input for the prognostic model.**

***Clinical Relevance*—This establishes an important prognostic model for the prediction of the trajectory of carotid artery atherosclerosis.**

I. Introduction

Stroke constitutes one of the most common and well-known causes of death in the EU with almost 450,000 casualties annually. However, with death being the ultimate outcome, stroke is also responsible for severe disabilities in adults. Over 50% of people that have suffered a stroke, end up depending totally on other people for simple, everyday activities. Besides the effect on the human factor, stroke affects global economy quite heavily, since the annual costs which are connected to stroke reach the enormous amount of a little less than €46 billion, including both direct and indirect costs, which account for the respective loss of productivity and health care [1, 2].

The advancement of carotid artery disease, which leads to larger atheromatic plaques that are susceptible to erosion or rupture, is one of the most important causes of stroke. When atheromatic plaques break, thromboembolism and cerebral infarction follow. It's interesting to note that more than 10% of all strokes are caused by thromboembolisms, which are caused by stenoses greater than 50% from asymptomatic plaques within the internal carotid artery (ICA). In cases of moderate to severe carotid artery disease, this increases the need for improved risk stratification methods that will improve patient care.

In this regard, numerous studies on simulating the development of atheromatic plaques and modeling the biological processes associated with atherosclerosis have been published [3-7]. Only a tiny number of these research studies have used patient-specific 3D artery models; the great majority of these are based on idealized 2D carotid models. Computed tomography (CT), magnetic resonance angiography (MRA), or ultrasound (US) scans can all be used to construct patient-specific 3D artery models. Then, using the 3D models of the lumen, outer wall, and plaque components, significant hemodynamic factors like endothelial shear stress (ESS), plaque structural stress (PSS), and areas of low ESS are calculated. These results are then used to simulate the infiltration of inflammatory cells and lipoproteins into the layers of the arterial wall, which is a process that promotes the progression of atheromatic plaque.

In the literature, various data-driven models have been presented for the risk stratification of CAD, with various attempts made to find the most significant biomarkers that are connected with the disease. Either statistical analysis or machine learning models serve as the foundation for these models. Greco *et al*. [8] created a model for the prediction of those at high risk for carotid degree of stenosis by utilizing

*This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 755320, as part of the TAXINOMISIS project.

P. K. Siogkas, D. S. Pleouras, V. D. Tsakanikas, M. D. Mantzaris, V. T. Potsika, A. Sakellarios, Evdokia Karamouzi and Foteini Lagiou are with the Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering University of Ioannina, Ioannina, Greece (e-mail: psiogkas4454@gmail.com,dipleouras@gmail.com,vasilistsakanikas@gmail.com,dmantzaris@gmail.com,potsika@gmail.com, ansakel13@gmail.com).

F. Sigala, George Charalampopoulos and George Galyfos are with the First Propaedeutic Dept. of Surgery, National and Kapodistrian University of Athens, Hippocration Hospital, Athens, Greece (e-mail: drfsigala@yahoo.gr, charalampopoulosg@gmail.com, georgegalyfos@hotmail.com).

Igor Koncar is with the Department of Vascular and Endovascular Surgery, Faculty of Medicine, University of Belgrade, Belgrade, Serbia (email: dr.koncar@gmail.com)

D. I. Fotiadis is with the Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering, University of Ioannina, and with the Biomedical Research Institute FORTH, University Campus of Ioannina, 45110 Ioannina, Greece (phone: +30 26510 09006; email: fotiadis@uoi.gr).

statistically based models. A CAD prediction model was established in a different study suggested by de Weerd *et al*. [9] to identify people with carotid artery stenosis greater than 50% and greater than 70% as well as to determine the most effective predictors of the carotid artery stenosis greater than 50% and greater than 70%. Finally, using high-resolution magnetic resonance imaging (MRI) modalities, Hao-wen Li *et al*. [10] conducted an observational study to separate the high-risk plaques (vulnerable plaques) and the low-risk (stable) carotid artery plaques.

In this work, we examine the efficacy of our proposed prognostic model regarding the progression of carotid artery stenoses, using fourteen heterogeneous features extracted from non-imaging, imaging and finite element blood flow simulation hemodynamic data.

II. Materials & Methods

A. Dataset

Twenty-seven patients with >50% carotid stenosis (forty-one carotid arteries) using a 1.5-T whole-body system (Signa HDx, GE Healthcare, Waukesha, WI, USA) with a bilateral four-channel phased-array carotid coil (Machnet BV, Eelde, the Netherlands) from the TAXINOMISIS cohort were used in the current study (Table 1). Patient provided written informed consent and enrolled in the TAXINOMISIS clinical study (www.clinicaltrials.gov; ID: NCT03495830) protocol which was approved by the local competent ethics committee.

Table 1: Demographics of the utilized dataset.

Patients (n=27)	**N (%)**
Age (years)	69.6±7.4
Gender (male)	21 (77.8)
Risk factors	
Smoking	20 (74)
Alcohol abuse	0
Diabetes mellitus	17 (62.9)
Hypertension	25 (92.6)
Hypercholesterolemia	27 (100)
Coronary disease	12 (44.4)
Obesity	3 (11.1)
BMI	25.1±6.9
Stenosis degree (%)	
50-70	29
70-90	10
90-100	2

B. 3D reconstruction

The ToF, T1w, T2w and PD series from the respective MRI screening were utilized by our in-house developed 3D reconstruction algorithm [11] to reconstruct in 3D the lumen of the 41 carotid arteries that were examined. By choosing to utilize only the lumen of each artery, we aimed to minimize the input needed for the blood flow simulations that were to be carried out.

C. Blood flow modeling

The 3D reconstructed luminal carotid geometries were then subjected to transient blood flow simulations, using patient-specific boundary conditions, as extracted from the respective carotid UltraSound (US) screening which includes flow velocity profiles for at least three consecutive cardiac cycles for each artery. Regarding the inlet, a patient-specific mass flow rate for an entire stabilized cardiac cycle was used as a boundary condition. The same process is followed for the external carotid artery (ECA), where a patient-specific mass flow rate profile is calculated and applied as an outlet boundary condition. The cardiac cycle duration is calculated from the patient-specific measured beats per minute and timesteps of 0.05 seconds are used to divide the cardiac cycle. Regarding the internal carotid artery (ICA), a zero-pressure boundary condition is applied as an outlet. A no-slip and no penetration boundary condition is applied at the arterial wall.

$$MAP = \frac{SBP + 2(DBP)}{3} \tag{1}$$

where MAP is the Mean Arterial Pressure, SBP the Systolic Blood Pressure and DBP the Diastolic Blood Pressure, respectively. The process described below was followed to calculate the pressure drop ratios between the CCA and the ICA or ECA outlets: using a zero-pressure boundary condition for the ICA outlet, we calculate the pressure gradient throughout the entire vessel. Having the patient-specific MAP value for the CCA which is calculated using Eq.1, we calculate the pressure difference for the simulated case between the CCA and the ECA and we subtract it from the patient-specific value, thus calculating the actual pressure for the ICA outlet. The same procedure is then applied between the CCA and the ECA and the final ECA outlet pressure value is calculated, thus allowing for the calculation of the two pressure drop ratios, respectively (i.e., P_{ECA}/P_{CCA} and P_{ICA}/P_{CCA}). To model blood flow in our simulations, we used the Navier-Stokes and the continuity equations:

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v} \bullet \nabla)\mathbf{v} - \nabla \bullet \boldsymbol{\tau} = 0, \tag{2}$$

$$\nabla \bullet (\rho \mathbf{v}) = 0, \tag{3}$$

where $\mathbf{v}$ is the blood velocity vector and $\boldsymbol{\tau}$ is the stress tensor, which is defined as:

$$\boldsymbol{\tau} = -p\delta_{ij} + 2\mu\varepsilon_{ij}, \tag{4}$$

where δ_{ij} is the Kronecker delta, μ is the blood dynamic viscosity, p is the blood pressure and ε_{ij} is the strain tensor calculated as:

$$\varepsilon_{ij} = \frac{1}{2}\left(\nabla \mathbf{v} + \nabla \mathbf{v}^T\right), \tag{5}$$

Blood was treated as Newtonian with density 1050 kg/m^3 and dynamic viscosity 0.0035 Pa·s, respectively. All simulations were carried out using ANSYS® v16.2. The element size was set to 0.16 mm or lower and constituted only of tetrahedra. The mesh size was determined after a thorough mesh sensitivity analysis. The convergence criterion was set to 10^{-4} and the iteration limit was 150 for each timestep.

D. Feature extraction

Fourteen (i.e., 14) features were extracted for the forty-one instances of our dataset. Time-averaged ESS values, total area of low ESS (i.e., <2 Pa), peak TAESS values, time averaged oscillatory shear index (OSI) values, areas of high OSI values, normalized areas of low TAESS, normalized areas of high OSI values, and P_{ECA}/P_{CCA} and P_{ICA}/P_{CCA} ratios were calculated for each arterial pair. OSI is calculated as:

$$OSI = \frac{1}{2}\left(1 - \frac{\int_0^T ESSdt}{\int_0^T ESSdt}\right) \quad (6)$$

where T is the cardiac cycle period used to calculate the time-averaged endothelial shear stress, and ESS is the endothelial shear stress. Furthermore, Time averaged Endothelial Shear Stress can be calculated through the following equation:

$$TAESS = \frac{1}{T}\int_0^T |ESS|\,dt \quad (7)$$

where T is the cardiac cycle period. The areas of low ESS values (i.e., TAESS<2 Pa) are prone to develop atheromatic plaque, or further progress the already existing one [12], whereas areas of high OSI values (i.e., OSI>0.35) are thought to be more likely to exhibit endothelial dysfunction and plaque development [13]. Table 2 depicts the utilized features and Fig.1 depicts the feature extraction workflow.

Table 2: Description of the feature dataset utilized as input.

Feature type	Feature
Imaging	Artery, Peak Systolic Velocity (PSV), ICA stenosis %, ECA stenosis %, Mean arterial pressure
Simulation-based	Peak TAESS, P_{ECA}/P_{CCA}, P_{ICA}/P_{CCA}, Vessel average TAESS, Vessel average OSI, Area of low TAESS, Area of high OSI, Normalized area of low TAESS, Normalized area of high OSI

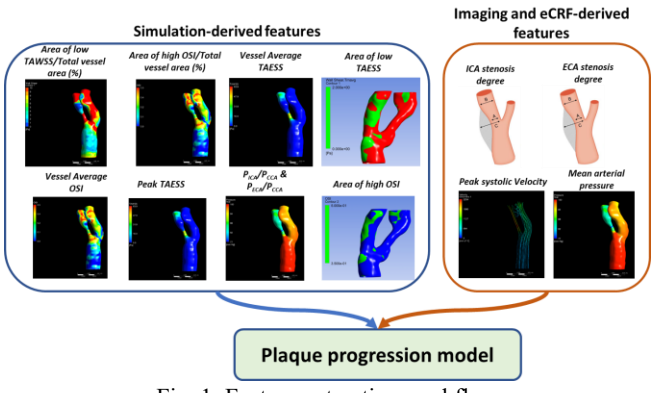

Fig. 1: Feature extraction workflow.

E. Problem definition

The carotid artery progression problem has been here formulated as a multivariate two class classification problem based on the stenosis geometrical findings deriving from the US screening of each patient. Elaborating on this, we analyzed the dataset, classifying arteries that exhibited no or minimal stenosis degree change as Class 0, and arteries that exhibited stenosis degree change higher than 10% as Class 1. We specifically chose 10% stenosis degree increase as a cut-off because percentages of 5% are not statistically significant and may be attributed to imaging artifacts or even miscalculation due to the very small difference in luminal diameter. Fig.2 depicts the illustration of the problem definition.

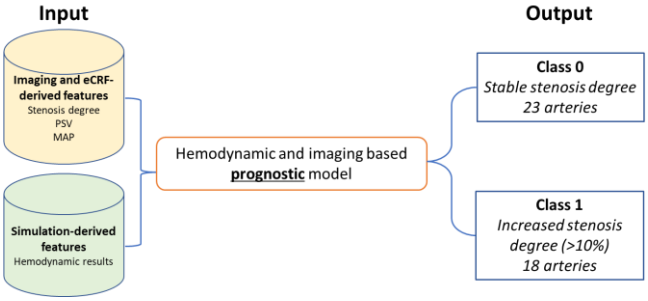

Fig. 2: Problem definition of the proposed machine learning based model.

F. Classification scheme implementation

Regarding the classification process, our input data are directly fed to five different classification algorithms. For the purposes of this work, we examine the efficacy of five well-known classification algorithms to distinguish the cases where the luminal diameter is prone to narrow by more than 10%. The classification algorithms that were used were Naïve Bayes, Random Forest, Support Vector Machine (SVM), Neural Network and Logistic Regression.

G. Model evaluation

A five-fold cross validation process was followed to evaluate our prognostic model, which divides the initial input dataset into four subsets, where the four subsets are used to train and the remaining one is used to test the algorithm, and after a full cross validation circle, the results of the five test sets are averaged to calculate the performance of the algorithm.

III. RESULTS

The classification results for all five classification algorithms are depicted in Table 3. Briefly, the best accuracy, area under the curve (AUC) and recall values were observed for the Neural Network classification scheme presenting with 0.707 accuracy and 0.725 AUC, respectively. The simple logistic regression scheme also presented with close values to the NN classifier with 0.698 AUC and 0.683 accuracy, respectively. Furthermore, increased stenosis degree by more than 10% (i.e., Class 1) is correlated with higher Gini decrease values with vessel average OSI, area of high OSI/total vessel area %, mean arterial pressure, area of low TAESS/total vessel area % and peak systolic velocity (PSV).

Table 3: Results obtained implementing different classification schemes.

Classifier	AUC	Accuracy	Precision	Recall
Neural Network	0.725	0.707	0.707	0.707
Logistic Regression	0.698	0.683	0.681	0.683
Naive Bayes	0.679	0.659	0.673	0.659
SVM	0.406	0.61	0.604	0.61
Random Forest	0.597	0.488	0.478	0.488

Fig. 3 depicts the confusion matrix for the NN classifier showing the proportion of predicted cases. The values in purple indicate the correctly predicted cases.

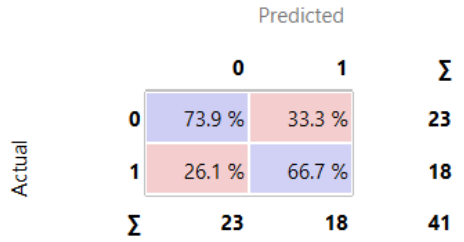

Fig. 3: Confusion matrix for the Neural Network classifier (showing proportion of predicted).

IV. DISCUSSION

In this proof-of-concept study, we presented a new prognostic model capable of predicting the progression of a carotid artery stenosis. Both imaging and non-imaging data were used as input for the model. Imaging data were utilized both for patient-specific flow profiles extraction for each artery, as well as, for the 3D reconstruction of the arterial lumen of the carotids that were used. Apart from the aforementioned data, several hemodynamic parameters that were calculated from respective patient-specific finite element-based blood flow simulations were also used as input for the prognostic model. The input features that were utilized were fourteen and the outcome of the model was either a stable luminal diameter, or an increase of the existing stenosis by at least 10%. Several classifier schemes were tested and the most efficient in terms of accuracy (i.e., 0.707) and area under the curve (i.e., AUC=0.725) was the neural network approach. Moreover, the most relevant features that affected the outcome the most were the metrics associated with areas of low TAESS and the OSI.

The main limitation of our study is the relatively small number of examined cases, however, even with such a low number of examined arteries, the prognostic capabilities of the proposed model were quite promising. The expansion of the dataset which is an ongoing process will enhance the capabilities of the model. We must also state that no data curation was needed for the utilized data since the quality of the data used was high and there were no missing values. Furthermore, the hemodynamic features that are used as input are highly dependent to the 3D reconstruction quality of the lumen, which in turn is also highly dependent from the quality of the MR images that are used.

V. CONCLUSION

In this work we proposed a prognostic model regarding the evolution of existing stenoses in the carotid vasculature, which is based on imaging, non-imaging and simulated hemodynamic features. Our methodology benefits from using machine learning models and provides a significant clinical impact regarding the choice of the type of medical treatment in carotid artery disease.

REFERENCES

[1] H. A. Wafa, C. D. A. Wolfe, E. Emmett, G. A. Roth, C. O. Johnson, and Y. Wang, "Burden of Stroke in Europe: Thirty-Year Projections of Incidence, Prevalence, Deaths, and Disability-Adjusted Life Years," *Stroke,* vol. 51, no. 8, pp. 2418-2427, Aug 2020, doi: 10.1161/STROKEAHA.120.029606.

[2] R. Naylor *et al.*, "Editor's Choice - European Society for Vascular Surgery (ESVS) 2023 Clinical Practice Guidelines on the Management of Atherosclerotic Carotid and Vertebral Artery Disease," *Eur J Vasc Endovasc Surg,* vol. 65, no. 1, pp. 7-111, Jan 2023, doi: 10.1016/j.ejvs.2022.04.011.

[3] D. S. Pleouras *et al.*, "Simulation of atherosclerotic plaque growth using computational biomechanics and patient-specific data," *Sci Rep,* vol. 10, no. 1, p. 17409, Oct 15 2020, doi: 10.1038/s41598-020-74583-y.

[4] A. I. Sakellarios *et al.*, "Natural History of Carotid Atherosclerosis in Relation to the Hemodynamic Environment," *Angiology,* vol. 68, no. 2, pp. 109-118, Feb 2017, doi: 10.1177/0003319716644138.

[5] N. A. Avgerinos and P. Neofytou, "Mathematical Modelling and Simulation of Atherosclerosis Formation and Progress: A Review," *Ann Biomed Eng,* vol. 47, no. 8, pp. 1764-1785, Aug 2019, doi: 10.1007/s10439-019-02268-3.

[6] A. Parton, V. McGilligan, M. O'Kane, F. R. Baldrick, and S. Watterson, "Computational modelling of atherosclerosis," *Brief Bioinform,* vol. 17, no. 4, pp. 562-75, Jul 2016, doi: 10.1093/bib/bbv081.

[7] A. J. Brown, Z. Teng, P. C. Evans, J. H. Gillard, H. Samady, and M. R. Bennett, "Role of biomechanical forces in the natural history of coronary atherosclerosis," *Nat Rev Cardiol,* vol. 13, no. 4, pp. 210-20, Apr 2016, doi: 10.1038/nrcardio.2015.203.

[8] G. Greco *et al.*, "A model for predicting the risk of carotid artery disease," *Ann Surg,* vol. 257, no. 6, pp. 1168-73, Jun 2013, doi: 10.1097/SLA.0b013e31827b9761.

[9] M. de Weerd *et al.*, "Prediction of asymptomatic carotid artery stenosis in the general population: identification of high-risk groups," *Stroke,* vol. 45, no. 8, pp. 2366-71, Aug 2014, doi: 10.1161/STROKEAHA.114.005145.

[10] H. W. Li *et al.*, "Association between ADAMTS7 polymorphism and carotid artery plaque vulnerability," *Medicine (Baltimore),* vol. 98, no. 43, p. e17438, Oct 2019, doi: 10.1097/MD.0000000000017438.

[11] V. D. Tsakanikas *et al.*, "A deep learning oriented method for automated 3D reconstruction of carotid arterial trees from MR imaging," *Annu Int Conf IEEE Eng Med Biol Soc,* vol. 2020, pp. 2408-2411, Jul 2020, doi: 10.1109/EMBC44109.2020.9176532.

[12] N. Sun, R. Torii, N. B. Wood, A. D. Hughes, S. A. Thom, and X. Y. Xu, "Computational modeling of LDL and albumin transport in an in vivo CT image-based human right coronary artery," *J Biomech Eng,* vol. 131, no. 2, p. 021003, Feb 2009, doi: 10.1115/1.3005161.

[13] D. N. Ku, D. P. Giddens, C. K. Zarins, and S. Glagov, "Pulsatile flow and atherosclerosis in the human carotid bifurcation. Positive correlation between plaque location and low oscillating shear stress," *Arteriosclerosis,* vol. 5, no. 3, pp. 293-302, May-Jun 1985, doi: 10.1161/01.atv.5.3.293.