Proceedings of the 3rd International Conference on Statistics: Theory and Applications (ICSTA'21)
~~Prague, Czech Republic~~ Virtual Conference – July, 2021
Paper No. 129
DOI: 10.11159/icsta21.129

Prediction Model for the Result of Percutaneous Coronary Intervention in Coronary Chronic Total Occlusions

Maria Ganopoulou[1], Ioannis Kangelidis[2], Georgios Sianos[3], Lefteris Angelis[4]
[1]School of Informatics, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece, marigano@csd.auth.gr;
[2]School of Medicine, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece,
johnykag@hotmail.com;
[3]School of Medicine, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece,
gsianos@auth.gr;
[4]School of Informatics, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece, lef@csd.auth.gr

Abstract - Coronary chronic total occlusions (CTOs) are very common in patients undergoing coronary angiography. There has been an increasing acceptance of the percutaneous coronary interventions (PCI) in CTOs. The success rate of PCI has been boosted over the last few years by, among else, operator experience and advances in technology, even achieving levels of approximately 90%. This study proposes a prediction model for the classification of the cases in successful and unsuccessful operations and addresses the problem of class imbalance in the response variable (operation result). It is based on the EuroCTO Registry, which is the largest database available worldwide consisting of 29,995 cases for the period 2008-2018. Binary logistic regression analysis and down-sampling were applied within a customized step-algorithm and standard statistical accuracy measures were employed for the assessment of the prediction model, such as sensitivity, specificity and the value of the area under the ROC (AUROC) curve. The analysis revealed new predictive factors, validating at the same time the impact of well-known predictors. A brief comparison has been performed with other models from the literature, which showed that the proposed model performs similarly or better than its contemporary competitors.

***Keywords*:** Coronary chronic total occlusions, down-sampling, logistic regression

Introduction

Coronary chronic total occlusions (CTOs) are defined as occlusions with duration at least 3 months with thrombolysis in myocardial infraction flow grade 0 (TIMI=0) [1]. Approximately 20% of the patients undergoing coronary angiography exhibit CTOs. These are the most complicated malfunctions of coronary arteries that an operator/cardiologist might address. The last decade, there has been an increasing acceptance of the percutaneous coronary interventions (PCI) in CTO, with increasing success rate [2]. There exist several attempts in the literature to assess the impact of different predictor variables to the success or failure of the PCI, ranging in size of the dataset from a few hundred cases [3] to approximately 20,000 [4]. In most studies, the authors develop a scoring system that categorizes patients in risk groups relevant to the success of the operation ([3]- [8]). The operation success ranges from 50% in the seminal paper of Morino et al. ([3]) to 93% in [6], and it depends on the experience of the operators and the time interval when data were collected, since advance in technology has a positive impact in the success of the operation. In this study, the analysis is based on the largest database available worldwide (EuroCTO Registry), which consists of 29,995 cases for the period 2008-2018 (164 variables including demographic data, clinical, anatomic, procedure parameters etc.). The aim of the study is to assess the impact of a large number of predictor variables in operation success, and to address for the first time the imbalance of the classes of the response variable (in this dataset it is 87.6% success and 12.4% failure).

Statistical Analysis

The aim of the statistical analysis was to develop an efficient predictive model for the operation success of the CTO PCI (response variable). For the purposes of the analysis, 32 out of the 164 variables were selected as predictor variables, based on conceptual, relevance, and literature criteria (e.g. technical parameters that are related with material, etc. were excluded). An additional criterion was the percentage of missing values, particularly variables with missing values over 20% were not assessed. No imputation methods for the missing values were considered, thus the analysis was based only on the observed data. The response variable is binary, and the association of the predictor variables with the response was assessed separately for each predictor variable with logistic regression analysis. Predictor variables with a p-value<0.05 (26 variables out of 32) were then included in a multivariate logistic regression model.

The following algorithm, which was based on [9], applied in order to assess the predictive power of our model:

1. Randomly draw 70% of the cases representing the training dataset.
2. Apply the full main effects model based on the training data.
3. Employ a stepwise approach to exclude variables from the full model based on the Akaike Information Criterion (AIC).
4. From the remaining variables, assess (with ANOVA) whether variables with corresponding p-value>0.05, or at least one corresponding p-value>0.05 (for categorical variables) can be excluded from the multivariable model.
5. For categorical variables assess whether merging of categories is possible.
6. Assess the model predictive accuracy on the test dataset (representing the remaining 30% of the original cases).

The assessment of the final model was based on specific statistical accuracy measures, i.e. sensitivity, specificity (the predicted probabilities for operation success were calculated for each case/patient, and specific thresholds were employed to categorize a case as operation success (predicted probability>=threshold) and as operation failure (<0.5, respectively)), and the value of the area under the ROC curve (AUROC).

To address the imbalance between the two classes of the response variable (12.4% Operation failure, 87.6% Operation success), a down-sampling approach was employed in the training data, resulting in a sample with equally -sized classes for the response variable (when down-sampling is used, it corresponds to an extra step (1a) in the step-algorithm above). The analysis was performed with R version 3.6.3.

Results

The data used after omitting any cases with at least one missing value (in the variables of interest) included 17,370 cases and 26 variables (including 25 predictors and the response). The proposed step-algorithm was initially applied without considering down-sampling, resulting in a model with 13 predictor variables, exhibiting very low specificity (prediction accuracy for the minority class - failure). To address this problem (probably due to the imbalance of the response classes), the step-algorithm was applied anew including down-sampling. The final model included 8 predictor variables. The results are displayed in Table 1, where apart from the standard threshold of 0.5, 4 more thresholds around 0.5 were considered to provide a short sensitivity analysis of the accuracy measures considered. Particularly, the sensitivity ranged from 49.39% to 81.57%, and the specificity from 47.22% to 80.71%. The AUROC was 0.7273.

Table 1: The values of sensitivity, specificity and AUROC in the test dataset for selected thresholds relevant to the predicted probabilities computed based on the logistic regression model including down-sampling.

	Prediction probabilities threshold				
Accuracy measures	0.4	0.45	0.5	0.55	0.6
Sensitivity	0.8157	0.7510	0.6725	0.5897	0.4939
Specificity	0.4722	0.5725	0.6636	0.7361	0.8071
AUROC	0.7273	0.7273	0.7273	0.7273	0.7273

The final model developed (based on the training data) included 8 main effects/predictors. The model coefficient estimates, and the corresponding Odds Ratio (with 95% confidence intervals) are displayed in Table 2. Most of these results are in agreement with the literature, e.g. it is known that the Proximal tortuosity classes Moderate and Severe are compounding regarding the success of the operation compared to the Straight class. Similarly the classes Moderate and

Severe of Calcification are compounding regarding the success of the operation compared to the None/Mild class. However, some new predictor variables were included in this model, such as Dyspnea and Bifurcation involvement.

In addition, the value 0.73 (p-value<0.001) of AUROC in the validation dataset is very good compared to the literature, since, in [4] the AUROC was 0.68, in [6] it was 0.72, and in [5] it was 0.68. In [3] the AUROC was found to be 0.76, however in this case, the data were far less (N=494), and the response classes were even.

Table 2: Coefficient estimates, and the corresponding Odds Ratio (with 95% confidence intervals) for the logistic regression model including down-sampling.

				95% CI for Exp(B)	
	B Estimate	p-value	Exp(B)	Lower	Upper
Constant	-.942	.001	.390		
Dyspnea_Yes	.294	<.001	1.342	1.144	1.575
Segmental [2] [2]wall motion CTO. Related lesion_H/A/D	-.408	<.001	.665	.569	.777
Bifurcation involvement_Yes	.308	.003	1.361	1.108	1.672
Visual estimation CTObrlength mm	-.015	<.001	.985	.981	.989
Visual estimation vessel diameter mm	.402	<.001	1.495	1.247	1.792
Proximal tortuosity		<.001			
Proximal tortuosity_Moderate	-.324	<.001	.723	.602	.868
Proximal tortuosity_Severe	-.993	<.001	.371	.277	.496
Calcification		<.001			
Calcification_Moderate (<=50% RLD)	-.411	<.001	.663	.552	.797
Calcification_Severe (>50% RLD)	-1.011	<.001	.364	.294	.450
Stump		<.001			
Stump_Tapered stump	1.246	<.001	3.475	2.755	4.383
Stump_Blunt stump	.671	<.001	1.956	1.561	2.452

Conclusion

This study has revealed the importance of factors that have not been included in similar modeling approaches in the literature to the best of our knowledge. This may indicate the need for further research in the field. On top of that, new methodology has to be introduced to address the imbalance in the classes of the operation result attributed to the high percentage of operation success, which is related, among else, with the increasing operator experience and advances in technology.

Acknowledgements

This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme 'Support for Researchers with Emphasis on Young Researchers' in the context of the project 'EDBM103-Part B, 2020-2022' (MIS: **5047860**).

References

[1] G. Sianos, G. Werner, A. Galassi , M. Papafaklis, J. Escaned, . D. Hildick-Smith, E. Christiansen, A. Gershlick, M. Carlino , . A. Karlas and N. Konstantinidis, "Recanalisation of chronic total coronary occlusions: 2012 consensus document from the EuroCTO club," *EuroIntervention,* vol. 8, no. 1, pp. 139-145, 2012.

[2] N. Konstantinidis, G. Werner, S. Deftereos, C. Di Mario, A. Galassi, J. Buettner, A. Avran, N. Reifart, O. Goktekin, R. Garbo, A. Bufe, K. Mashayekhi, N. Boudou, M. Meyer-Geβner, B. Lauer, S. Elhadad, E. H. Christiansen, J. Escaned, D. Hildick-Smith, M. Carlino, Y. Louvard, T. Lefèvre, L. Angelis, G. Giannopoulos and G. Sianos, "Temporal trends in chronic Total occlusion interventions in Europe: 17 626 procedures from the European registry of chronic Total occlusion," Circulation: Cardiovascular Interventions, vol. 11, no. 10, p. e006229, 2018.

[3] Y. Morino , M. Abe , T. Morimoto , T. Kimura , Y. Hayashi , T. Muramatsu , M. Ochiai , Y. Noguchi , K. Kato , Y. Shibata and Y. Hiasa , "Predicting successful guidewire crossing through chronic total occlusion of native coronary lesions within 30 minutes: the J-CTO (Multicenter CTO Registry in Japan) score as a difficulty grading and time assessment tool," *JACC: Cardiovascular Interventions,* vol. 4, no. 2, pp. 213-221, 2011.

[4] Z. Szijgyarto , R. Rampat , G. Werner , C. Ho , N. Reifart , T. Lefevre , Y. Louvard , A. Avran , M. Kambis , H. Buettner and C. Di Mario , "Derivation and validation of a chronic total coronary occlusion intervention procedural success score from the 20,000-patient EuroCTO registry: the EuroCTO (CASTLE) score," *JACC: Cardiovascular Interventions,* vol. 12, no. 4, pp. 335-342, 2019.

[5] G. Alessandrino , B. Chevalier , T. Lefèvre , F. Sanguineti , P. Garot , T. Unterseeh , T. Hovasse , M. Morice and Y. Louvard , "A clinical and angiographic scoring system to predict the probability of successful first-attempt percutaneous coronary intervention in patients with total chronic coronary occlusion," *JACC: Cardiovascular Interventions,* vol. 8, no. 12, pp. 1540-1548, 2015.

[6] G. Christopoulos , D. Kandzari , R. Yeh, F. Jaffer , D. Karmpaliotis , M. Wyman , K. Alaswad , W. Lombardi , J. Grantham, J. Moses and G. Christakopoulos , "Development and validation of a novel scoring system for predicting technical success of chronic total occlusion percutaneous coronary interventions the PROGRESS CTO (Prospective Global Registry for the Study of Chronic Total Occlusion Intervention) score," *JACC: Cardiovascular Interventions,* vol. 9, no. 1, pp. 1-9, 2016.

[7] J. Maeremans , J. Spratt , P. Knaapen , S. Walsh , P. Agostoni , W. Wilson , A. Avran , B. Faurie , E. Bressollette , P. Kayaert and A. Bagnall , "Towards a contemporary, comprehensive scoring system for determining technical outcomes of hybrid percutaneous chronic total occlusion treatment: the RECHARGE score," *Catheterization and Cardiovascular Interventions,* vol. 2, no. 91, pp. 192-202, 2018.

[8] A. Galassi , M. Boukhris , S. Azzarelli , M. Castaing , F. Marzà and S. Tomasello , "Percutaneous coronary revascularization for chronic total occlusions: a novel predictive score of technical failure using advanced technologies," *JACC: Cardiovascular Interventions,* vol. 9, no. 9, pp. 911-922, 2016.

[9] M. J. Crawley, The R book, John Wiley & Sons, 2012.