

A Study on Predicting Cause of Side Effects in ACS Patients Using Machine Learning

Sang Hyun Park¹, Eun Soo Choi², Min Soo Kang^{3*}

¹Department of Internal Medicine, Eulji Hospital, Korea, psh@eulji.ac.kr

² Department of Medical IT Marketing, Eulji University, Korea, ces951030@gmail.com

³Department of Medical IT, Eulji University, Korea, mskang@eulji.ac.kr (*corresponding author)

Article Info

Volume 81

Page Number: 2255 - 2261

Publication Issue:

November-December 2019

Abstract

Cardiovascular disorders are the 1st ranked cause of death throughout the world and, in particular, ACS (Acute Coronary Syndrome) is the key risk factor of cardiovascular disorders. ACS is a fatal disorder with high risk of manifestation of MI (Myocardial Infarction) and death rate that imparts significant burden on individuals and society due to high cost incurred in treatment. Clopidogrel or Ticagrelor is used as anti-platelet drug. Recently, Ticagrelor known to be high safe and effective has been administered predominantly. However, it has become difficult to administer this drug because of symptoms of dyspnea induced in some patients. Accordingly, statistically significant attributes were sought with Pearson correlation technique to assess the cause of dyspnea in patients and the causes of dyspnea in patients were predicted by applying Two-class Decision Forest algorithm. As a result, it was found that 7 attributes including Uric acid, CPK, HDL cholesterol, TSH, Age, AST and LA Volume index are related to dyspnea, and in particular, Uric Acid imparted greatest effect on dyspnea. It would be possible to minimize the side effects by screening administration of Ticagrelor to ACS patients by utilizing such results of research.

Article History

Article Received: 5 March 2019

Revised: 18 May 2019

Accepted: 24 September 2019

Publication: 12 December 2019

Keywords: ACS, Dyspnea, Machine Learning, Ticagrelor, Uric Acid

1. INTRODUCTION

Cardiovascular disorders are the 1st ranked cause of death throughout the world and ACS (Acute Coronary Syndrome) is the key risk factor of cardiovascular disorders. Acute coronary syndrome is the term used to refer to all acute symptoms that arise from the degradation in the ability to supply blood to cardiac muscles and is a fatal disorder with high risk of manifestation of MI (Myocardial Infarction) and death rate in patients. Moreover, it imparts significant burden to individuals and society due to high cost incurred in treatment. In spite of the advancement of treatment drugs for ACS, death rate due to atherothrombotic event that follows diagnosis of

ACS is high, thereby requiring assertive treatments following ACS diagnosis. There are various methods for these treatments, which fundamentally remove or suppress the progress of thrombus. Clopidogrel or Ticagrelor is used to either remove or suppress formation of thrombus [1]. Anti-platelet effects of thienopyridine range drugs such as Clopidogrel, etc. have characteristic of irreversibility, thereby requiring generation of new platelet to terminate its anti-platelet effect. Therefore, it can become a problem in the event of emergency surgery with lack of time required for generation of platelet. For this reason, there is a need for anti-platelet drug that achieves more outstanding IPA (Inhibition of platelet aggregation) than

Clopidogrel with characteristics including quick commencement and termination of drug actions, no need for metabolic activation process in the body, more consistent effects among patients and no increase in the risk of hemorrhage.

Ticagrelor is the first drug of CPTP (Cyclopentyl- Triazolopyrimidine), which is a new chemical range of anti-platelet drugs, and has the characteristics that are importantly distinguished from those of thienopyridine. Accordingly, Ticagrelor is readily absorbed and displays its effectiveness quickly after oral administration, and promptly reaches more consistent and higher IPA than Clopidogrel [2]. Since this process does not need generation of new platelet or injection of platelet, pharmacological and clinical profiles of Ticagrelor demonstrate that Ticagrelor can provide more consistent, quicker and more effective anti-platelet effects than Clopidogrel or Prasugrel. Therefore, Ticagrelor is administered to ACS patient for whom the drug is relatively safe with quick effectiveness. Although the symptoms of dyspnea are manifested in some of the patients [3], evidence of cardiopulmonary dysfunction was not detected [4]. In addition, it was found through post-hoc analysis of PLATO research that the extent of dyspnea is in the range of mild to moderate, and it was demonstrated that it does not impart influence on the efficacy or safety results of ACS patient [5]. According to the Barbieri, L., et al. "Uric acid and high-residual platelet reactivity in patients treated with Clopidogrel or Ticagrelor. Nutrition, Metabolism and Cardiovascular Diseases" in 2016, uric acid level does not induce reaction against Ticagrelor and Clopidogrel or ASA (Adams-Stokes attack) [6]. However, these that focused on the correlation between Ticagrelor and dyspnea asserted that it was discovered through PLATO clinical study that Ticagrelor is associated with dyspnea and bradyarrhythmias [7][8]. M.Gaubert et al. reported that such

clinical results can lead to cessation of administration of Ticagrelor to ACS patient [9]. According to "Zhang, N., Zhang, Z., Yang, Y., Xu, Y., Li, G., & Liu, T. Ticagrelor-related gout: An underestimated side effect. International journal of cardiology" in 2015, precautions need to be taken periodically for the serum uric acid levels when administering Ticagrelor and, in particular, greater precautions need to be exercised in the case of hyperuricemia patient [10].

Dyspnea is a side effect because it inhibits the use of Ticagrelor and cannot effectively treat patients with ACS. To predict these side effects, we used a machine learning based decision tree method.

Machine learning is an area of artificial intelligence that develops algorithm and technology that enables computer to study with focus on representation and generalization [11]. The term machine learning was first used by Arthur Samuel, who was a researcher in the area of artificial intelligence at IBM, in his thesis, "Studies in Machine Learning Using the Game of Checkers." Here, machine refers to programmable computer and is also commonly refers to the server [12]. Machine learning has been researched through 3 approach methods. First is the nerve model paradigm. Nerve model began with perceptron and has now progressed into deep learning. Second is the study paradigm under the concept of symbols. This paradigm uses logic or graph structure rather than number or statistical theory, and was the core approach method of artificial intelligence from the mid-1970's to latter part of 1980's. Third is the intensive modern knowledge paradigm. This paradigm that began from the middle of the 1970's by subletting the nerve model that begins study from the state of blank slate and with the emergence of the theory that already learned knowledge need to be reutilized [13]. Machine learning is divided into supervised learning and unsupervised learning. In this study, decision tree

method, which is an area of supervised learning [14].

Decision tree can also be used in classification or regression. Decision tree was selected as the method for dyspnea of patient administered with Ticagrelor since it can be used in finding regularities and when predicting the future event by using such regularities in the event of assessing to which group a given data belongs to by discovering the characteristics for each group through division into several groups with similar characteristics.

2. SUPERVISED LEARNING

Machine learning of supervised learning has several approach methods in accordance with the characteristics and purposes of the data set, and supervised learning is one such method. Supervised learning is a learning model to find the which resultant value new data has, after having created determination equation for distinction of the data through various algorithm with labeled learning data set [15]. Supervised learning is divided largely into regression and classification with regression analysis being the representative of supervised learning. Since both classification and regression are supervised learning models, they have the common feature of pursuing learning with labeled input data. In terms of the difference between classification and regression, classification has fixated resultant value while the resultant value of regression can be any value within the range of the data set [16]. More specifically, classification model has the goal of finding the group to which the newly inputted data belongs to after having learned with learning data to which the group name (label) has been attached to. Therefore, the resultant value of classification model must be one of the labels of the learning data. For example, if there is a data set composed of labels including diabetes, age, height, blood pressure and weight, the resultant value of classification model must be one of

diabetes, age, height, blood pressure and weight.

2.1 Classification

Dictionary meaning of classification refers to the process of cognition, differentiation and understanding the concept or principal, with the implication that the principal is within the category and generally has specific purpose. Therefore, it is essential in all types of interaction related to language, presumption, inference, decision making and environment. Classification in machine learning is a representative technique of data analysis that classifies object with multiple numbers of attributes or variables into one of the groups or classes that had been determined in advance [17]. Classification model is divided into various models in accordance with algorithm. Among these, K-NN (K-Nearest Neighbor), SVM (Support Vector Machine) and decision tree are the representative models. This study utilized decision tree method among the classification algorithms.

2.2 Decision Tree

Fundamentally, Decision Tree has the following advantages in classification in general [18]. It is possible to discern the non-linear decision boundaries. It is efficient in computation and memory usage. Feature selection is integrated in the model learning and classification process. Model can be made by applying noise and numerous features. It can handle data with varied distributions.

This study also sought to find the attributes that affect the dyspnea of patient administered with Ticagrelor by utilizing the advantages of the feature selection of Decision Tree.

3. EXPERIMENT

3.1 Experimental Environment

In this paper, Microsoft Azure Machine Learning Studio was used to assess the dyspnea attribute and to find the most influential attribute when

Ticagrelor is administered to ACS patient. Azure is a cloud computing platform of Microsoft that began servicing since 2010 along with the commencement of PaaS service in 2011 followed by IaaS service in 2013. Azure platform provides more than 600 services and we are using Azure Machine Learning Studio, which is one of such services [19]. Azure Machine Learning Studio is used because of the provision of modules in various formats that assist with the development of and distribution of one's machine learning model in actual environment. Moreover, it is void of the lack of flexibility in the computing resources necessary and inconvenience of the setting works for GPU-based learning such as Tensor Flow library, which are the chronic problem of the existing Machine Learning library and tool, thereby making it very easy to install and set tools and environment necessary for learning [20]. In addition, it solves the difficulties of recording of experimental processes and version management, and enables easy direct installation and setting of solution system for collaboration with Jupyter Notebook, etc. Utilization of Azure Machine Learning Studio enables quick execution of repetitive works by producing and testing several models within several minutes. Generally, designing of prediction model, and editing and experimenting of parameter or model are executed repetitively when producing experiment. Series of such works can be handled easily through the utilization of Azure Machine Learning Studio.

3.2 Data Preprocessing

This study utilized the clinical data of the Eulji Hospital of Eulji University. Use of the data for this study was approved through deliberation by the Institutional Review Board of Eulji University (IRB No. EMC 2018-01-005).

Cardiovascular disorders are the 1st ranked cause of the death throughout the world and, in particular, acute coronary syndrome (ACS) is the

key risk factor of cardiovascular disorders. Blood test and EMR/OCS data were collected from ACS patients, in whom symptoms of dyspnea were manifested following administration of Ticagrelor, as the subjects. Pre-processing of data collected was executed. Figure 1 illustrates the pre-processing procedure for data.

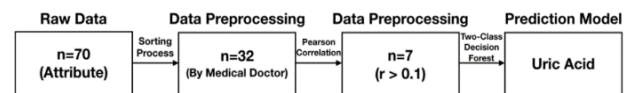


Figure 1 : Data pre-processing procedure

Clinical data of 166 patients for whom Ticagrelor was administered, including 22 patients with dyspnea and 144 patients without dyspnea, at the Eulji Hospital of Eulji University were utilized. These ACS patients participated in the clinical experiment conducted at the Eulji Hospital of Eulji University. If administration of Ticagrelor resulted in continuous efficacy, the treatment was maintained continuously for the patient while administration of Ticagrelor was stopped for patient who displayed symptoms of dyspnea within prescribed period of time. There is a need for pre-processing stage of anonymity for the data collected from 166 patients administered with Ticagrelor, and categorizing the 70 attributes of attributes of the drug administered patient into those that can be learned and those that cannot be learned. 32 attributes were screened out as being learnable through the Sorting Process by medical specialist. Then, data among the selected attributes that were not digitized were defined as follows for learning.

Definition 1: Attribute values that were not measured in the clinical process were not uses as attributes for training.

Definition 2: When the attribute value measured through clinical setting exceeded the maximum value, the maximum value was applied.

Out of the 32 attributes that were finally selected, 7 attributes were sorted out through Pearson

correlation technique to find the attributes that affect patients with dyspnea. As the result, 7 attributes including uric acid, CPK, HDL cholesterol, TSH, Age, AST and LA Volume iex were detected. Two-Class Decision Forest algorithm in 8 levels was applied to find the most influential attribute among the 7 attributes detected [21].

3.3 Experiment Result

In order to find the attributes that affect patients with dyspnea, Pearson correlation technique was used. As the results, 7 attributes including uric acid, CPK, HDL cholesterol, TSH, Age, AST and LA Volume iex were found, through the results, to affect dyspnea.

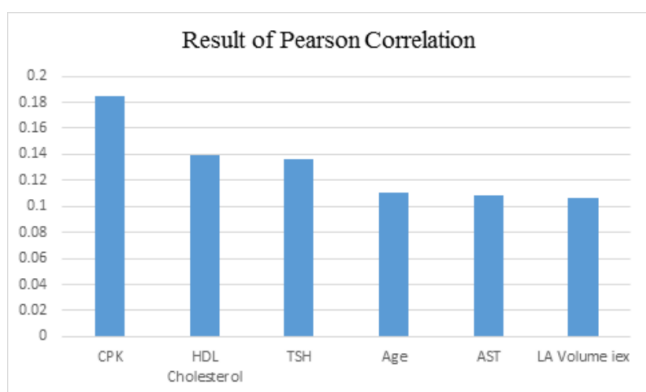


Figure 2 : Pearson Correlation Results

Figure 2 illustrates Pearson Correlation Results. Pearson correlation coefficient is used universally to compute the relevance between 2 variables.

Coefficient of determination is computed as, which signify the extent of predictability of from [22]. If is in the range of +0.1 ~ +0.3, there is weak positive linear relationship, and the aforementioned 7 attributes, that is, uric acid, CPK, HDL cholesterol, TSH, Age, AST and LA Volume iex, can be deemed to have relationship with dyspnea.

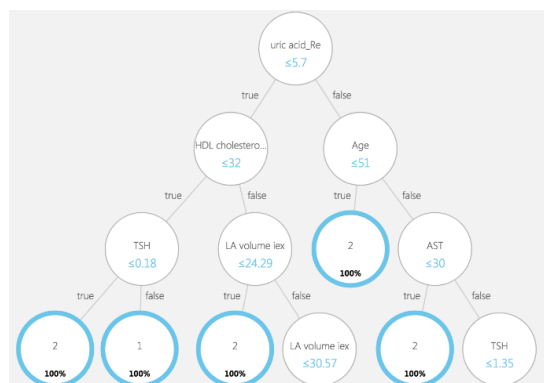


Figure 3 Top position tree

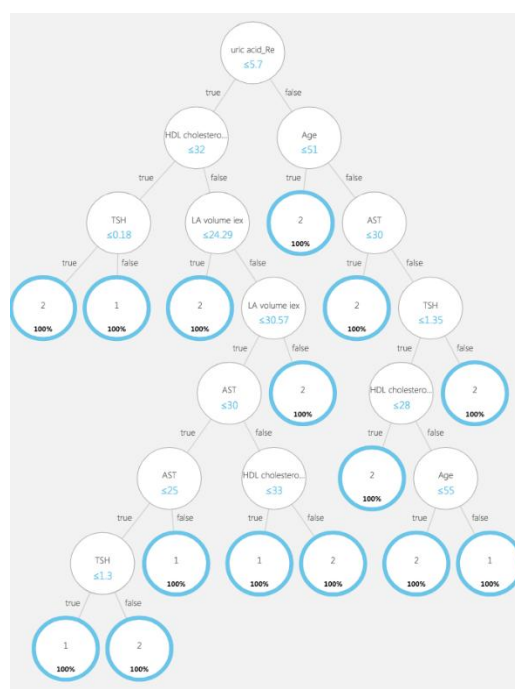


Figure 4 : Entire tree

Figure 3 and Figure 4 illustrates the results of having applied Two-Class Decision Forest by selecting the level of dyspnea on the basis of the 7 main attributes selected above. Figure 4 illustrates the entire tree of the learned Two-Class Decision Forest while figure 3 illustrates the enlargement of the top portion of figure 4 [23-25].

It can be confirmed that uric acid, through the high-ranking determinant of figure 3, has influence on the manifestation of dyspnea through administration of Ticagrelor.

4. CONCLUSION

In this study, 7 attributes significant for the dyspnea was sought through Pearson correlation technique in order to assess the cause of the dyspnea of ACS patient when administered with Ticagrelor. As the result, uric acid, CPK, HDL cholesterol, TSH, Age, AST and LA Volume iex were found to be the attributes imparting effect on dyspnea. In particular, value of uric acid is found to have the greatest significance in relation to dyspnea. Moreover, it was confirmed that uric acid is an attribute with influence on the manifestation of dyspnea through the trained tree by applying the 7 attributes to Two-Class Decision Forest algorithm.

Finding that uric acid is the attribute with the greatest influence on ACS patient administered with Ticagrelor by using machine learning was a result similar to the clinical results. It would be possible to minimize the side effects if administration of Ticagrelor to ACS patient is determined by utilizing the researched results. Moreover, utilization of more extensive range of data will further elevate the level of utilization of machine learning throughout the entire range of medical field through elevation of the accuracy of machine learning.

ACKNOWLEDGEMENT

This paper was supported by Eulji University in 2019

REFERENCES

1. Doo Soo Jeon, **Diagnosis and management of acute coronary syndrome**, Journal of the Korean Medical Association, Vol.60, No.7, 2017, pp. 568–576.
2. L. Wallentin, R.C. Becker, A. Budaj, C.P. Cannon, H. Emanu-elsson, C. Held, et al., **Ticagrelor versus clopidogrel in patients with acute coronary syndromes**, N. Engl. J. Med. Vol.361, 2009, pp 1045-1057.
3. Chul Min Kim, **Acute coronary syndrome**, Spring Conference, pp 97-100, KJHP, 2013
4. Storey R. F., Bliden K. P., Patil S. B., et al. **Incidence of dyspnea and assessment of cardiac and pulmonary function in patients with stable coronary artery disease receiving ticagrelor, clopidogrel, or placebo in the ONSET/OFFSET study**, Journal of the American College of Cardiology, Vol.56, No.3, 2010, pp. 185-193.
5. Storey R. F., Becker, R. C., Harrington, R. A., Husted, S., James, S. K., Cools, F., ... & Cairns, R., **Characteriza-tion of dyspnoea in PLATO study patients treated with ticagre-lor or clopidogrel and its association with clinical out-comes**, European heart journal, Vol.32, No.23, 2011, pp 2945-2953.
6. Barbieri, L., Verdoia, M., Pergolini, P., Nardin, M., Rolla, R., Marino, P., ... & Novara Atherosclerosis Study Group., **Uric acid and high-residual platelet reactivity in patients treated with clopidogrel or ticagrelor. Nutrition, Metabolism and Car-diovascular Diseases**, Vol.26 No.4, 2016, pp 352-358.
7. M.J. Sanchez-Galian, P.J. Flores-Blanco, A. Lopez-Cuenca, M. Gomez-Molina, E. Guerrero-Perez, F. Cambronero-Sanchez, et al., **Ticagrelor related dyspnea in patients with acute coronary syndromes: incidence and implication on ticagrelor withdrawn**, Int. J. Cardiol. Vol.187, 2015, pp 517-518.
8. N. Lombardi, M.C. Lenti, R. Matucci, A. Mugelli, A. Vannacci, **Ticagrelor-related dyspnea: an underestimated and poorly managed event?**, Int. J. Cardiol. Vol. 179, 2015, pp 238-239.
9. M.Gaubert, M.Laine, T.Richard, N.Fournier, C.Gramond, J.Bessereau et al., **Effect of ticagrelor-related dyspnea on compliance with therapy in acute coronary syndrome patients**, Int. J. Car-diol. Vol. 173, 2014, pp 120-121.
10. Zhang, N., Zhang, Z., Yang, Y., Xu, Y., Li, G., & Liu, T., **Ticagrelor-related gout: An**

- underestimated side effect. International journal of cardiology**, Vol.192, 2015, pp 11-13.
11. Rafols, E. J., Ring, M. B., Sutton, R. S., & Tanner, B., **Using Predictive Representations to Improve Generalization in Reinforcement Learning**, In IJCAI, 2005, pp. 835-840.
 12. Samuel, Arthur L. **Some studies in machine learning using the game of checkers**, IBM Journal of research and development 3.3, 1959, pp 210-229.
 13. Eui Joong Kim, **Artificial Intelligence, Machine Learning, Deep Learning introduction**, wikibooks, 2016
 14. Friedman, J., Hastie, T., & Tibshirani, R., **The elements of statistical learning**, Vol. 1, No. 10. New York, NY, USA:: Springer series in statistics, 2001
 15. Marsland, Stephen. **Machine learning: an algorithmic perspective**. Chapman and Hall/CRC, 2011.
 16. Smola, A., & Vishwanathan, S. V. N., **Introduction to machine learning**. Cambridge University, UK, 2008, pp 32-34
 17. James, Gareth, et al. **An introduction to statistical learning**. Vol. 112. New York: springer, 2013.
 18. Criminisi, A., Shotton, J., & Konukoglu, E., **Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning**, Microsoft Research, 2011
 19. Microsoft Azure Machine Learning, Website : <https://docs.microsoft.com/ko-kr/azure/machine-learning/studio/what-is-ml-studio>
 20. Min Soo Kang, Eun Soo Choi, **Getting started Machine Learning with MicroSoft AZURE ML**, Hanti Media, 2018
 21. Microsoft Azure Machine Learning, Website : <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/two-class-decision-forest>
 22. Asuero, A. G., Sayago, A., & Gonzalez, A. G., **The correlation coefficient: An overview. Critical reviews in analytical chemistry**, Vol. 36, No.1, 2006, pp 41-59.
 23. Petrenko, Y.; Vechkinzova, E.; Antonov, V. 2019. Transition from the industrial clusters to the smart specialization of the regions in Kazakhstan, Insights into Regional Development 1(2): 118-128. [https://doi.org/10.9770/ird.2019.1.2\(3\)](https://doi.org/10.9770/ird.2019.1.2(3))
 24. Mayorova, E. 2019. Corporate social responsibility disclosure: evidence from the European retail sector. Entrepreneurship and Sustainability Issues, 7(2), 891- 905. [http://doi.org/10.9770/jesi.2019.7.2\(7\)](http://doi.org/10.9770/jesi.2019.7.2(7))
 25. Jabarullah, N.H. (2019) Production of olefins from syngas over Al₂O₃ supported Ni and Cu nano-catalysts, Petroleum Science and Technology, 37 (4), 382 – 385.