KSN SEOUL, KOREA 2023
April 27(Thu) - 30(Sun), 2023
Coex, Seoul, Korea

Abstract Type : Poster
Abstract Submission No. : 1442

Erythropoiesis stimulating agent recommendation model using artificial intelligence in patients with chronic kidney disease

Byeongo Choi, Yoo jin Lee, Chang min Heo, Sihyung Park, Bong Soo Park, Yeongrok Oh, Jiyae Yi, Yang Wook Kim
Department of Internal Medicine-Nephrology, Inje University Haeundae Paik Hospital, Korea, Republic of

Objectives: Anemia in chronic kidney disease is caused by lack of erythropoietin. Treatment is performed by administering erythropoiesis-stimulating agents, which is determined based on empirical predictions, but some patients fall below or exceed the target. We tried to estimate the appropriate dose of the erythropoiesis-stimulating agents using a machine learning algorithm.

Methods: A machine learning model has been developed to predict the appropriate dose of darbepoetin alfa for a total of 125 patients. The linear regression model used both a complete model with all variables and a reduced model with variable selection applied by regression diagnostics and multi-collinear removal. The machine learning model used random forest and XGboost models.

Results:
The linear regression model has a coefficient of determination of 0.818 and the estimated regression formula can account for 81.8% of the total erythropoiesis stimulating agent dose. The only variable affecting single dose of erythropoiesis stimulating agent was the level of hemoglobin before administration. We compared the full and reduced models of linear regression analysis and the predictive performance of the random forest and XGboost models of machine learning with the absolute average error of the test dataset and each model. In linear regression analysis, the full model was 8.889 and the reduced model was 12.217, indicating that the full model showed relatively high predictive power. In machine learning, the random forest is 5.474 and the XGboost is 3.952, indicating that the XGboost model showed relatively high predictive power. The complete model of the linear regression model did not meet the regression assumptions such as multi-collinearity and residual independence, but performed better than the reduced model only in terms of prediction rather than interpretation.

Conclusions: We need to flexibly utilize machine learning techniques that are advantageous for predicting detection power and statistical analysis techniques that are advantageous for interpreting results.