Iran J Public Health, Vol. 51, No.4, Apr 2022, pp.964-966

Letter to the Editor

An Artificial Neural Network Approach to Predicting Stroke in Postmenopausal Women

Hyejin Park [1], ***Kisok Kim*** [2]

1. *Department of Health Sciences, Dongduk Women's University, Seoul 02748, Republic of Korea*
2. *College of Pharmacy, Keimyung University, Daegu 42601, Republic of Korea*

***Corresponding Author:** Email: kimkisok@kmu.ac.kr

(Received 10 Oct 2020; accepted 21 Oct 2020)

Dear Editor-in-Chief

Stroke is a leading cause of disability and death among elderly women in developed countries (1). Because the incidence of stroke dramatically increases after menopause, most women experience stroke relatively late in life. Furthermore, the stroke incidence is particularly high in Asians and blacks (1, 2). In 2017, the prevalence of stroke in elderly South Korean women was 5.1% (3).

Although menopause is a pathophysiological change that begins at an average age of 50 years, the mechanism by which menopause exerts its effects on vascular incidents remains unknown (4). Although the majority of strokes can be attributed to risk factors common to both men and women, including high blood pressure, high cholesterol levels, diabetes, smoking, and overweight status, there are a number of stroke risk factors that are specific to women, including those that involve sex hormones and exogenous estrogens (5).

Artificial neural network (ANN) prediction models are now used in many areas of healthcare research (6). ANN models are ideal for predicting disease prevalence in individuals, as they involve calculating nonlinear correlations between independent and outcome variables until a high level of accuracy is achieved (7).

Therefore, we aimed to develop an effective ANN model for predicting stroke in postmenopausal women based on data from the Korea National Health and Nutrition Examination Survey (KNHANES).

We extracted data, including age, income, educational attainment, physical exercise, body mass index, waist circumference, smoking and drinking habits, and food intake, from the KNHANES 2010–2015 databases. Daily energy intake was assessed via 24-h recall and food-intake frequency methods. The data on all variables were then transformed into values ranging from 0 to 1 using the MinMaxScalar class from the preprocessing module of the Python scikit-learn library. After data processing, the entire data set was split into training and test sets at a ratio of approximately 7:3. The ANN model consisted of an input layer of 9 dimensions, hidden layer of 12 dimensions, and output layer of 1 dimension. The network was trained using stochastic gradient descent and optimized using Adam optimizer with a learning rate of 0.001. The neural network was trained for 100 epochs. We used the Rectifed Linear Unit activation function at each layer and the Sigmoid at the output layer. A dropout regularization of 30% was applied at the input layer, and the categorical cross entropy error function

for binary classification was used as the loss function (Fig. 1). The ANN model was implemented using the Tensorfow framework (version 1.12.0). We calculated the accuracy of the test set to measure the performance of the prediction model. Probability estimates were calculated to plot a receiving operating characteristic (ROC) curve for the model, and the ROC curve was used to measure the predictive performance of the model.

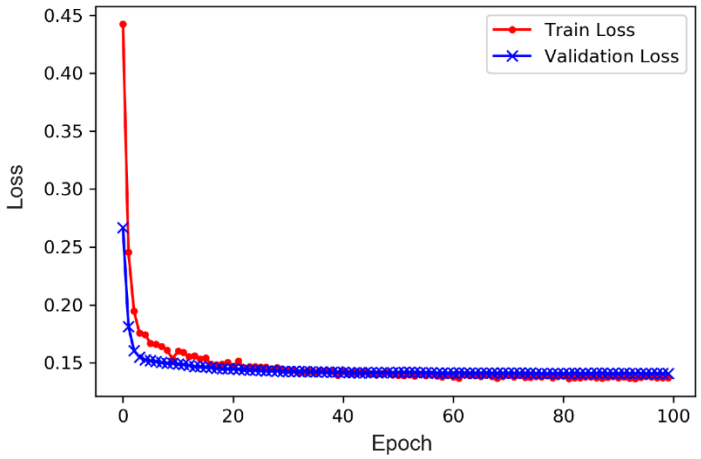

Fig. 1: Plot of train and validation loss as a function of epoch

The mean age of the 8524 participants included in the study was 64.2 years, and the stroke prevalence was 3.4%. After normalization of all variables, the ANN model achieved a relatively high accuracy rate of 0.966. The area under the ROC curve for the ANN prediction models was 0.689 (95% confidence interval = 0.626–0.752) (Fig. 2).

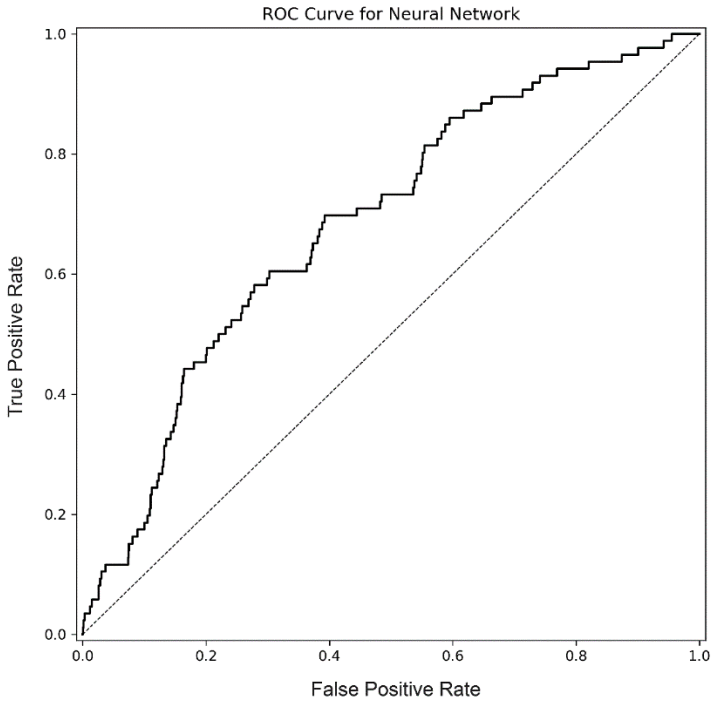

Fig. 2: Receiver operating characteristics (ROC) curves for artificial neural network in prediction of stroke

ANN models have recently been used to predict the occurrence of many diseases, including hypertension, diabetes, and heart failure (8–10). In this study, we developed a high-performance stroke prediction ANN model based on sociodemographic variables. This prediction model, based on nationwide survey data, can be used as a preliminary screening tool to identify postmenopausal women at high risk of stroke. The ANN-based prediction model may increase the chance of early intervention and reduce treatment costs.

Acknowledgements

This work was supported by research grants from the Daegu Catholic University in 2018 (No. 20181198).

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Bushnell CD, Hurn P, Colton C, et al (2006). Advancing the study of stroke in women: summary and recommendations for future research from an NINDS-sponsored multidisciplinary working group. *Stroke*, 37(9): 2387-99.
2. Atsma F, Bartelink ML, Grobbee DE, van der Schouw YT (2006). Postmenopausal status and early menopause as independent risk factors for cardiovascular disease: a meta-analysis. *Menopause*, 13(2): 265-79.
3. Korean Statistical Information Service (2018). State of major illnesses and symptoms by province, grade, qualification, 2017. National Statistical Office, Republic of Korea. www.koisis.kr
4. Hu FB, Grodstein F, Hennekens CH, Colditz GA, Johnson M, Manson JE, Rosner B, Stampfer MJ (1999). Age at natural menopause and risk of cardiovascular disease. *Arch Intern Med*, 159(10): 1061-6.
5. National Stroke Association (2019). Women and stroke. American Heart Association, USA. www.stroke.org
6. LeCun Y, Bengio Y, Hinton G (2015). Deep learning. *Nature*, 521(7553): 436-44.
7. Li H, Luo M, Zheng J, Luo J, Zeng R, Feng N, Du Q, Fang J (2017). An artificial neural network prediction model of congenital heart disease based on risk factors: A hospital-based case-control study. *Medicine*, 96(6): e6090.
8. Huang SQ, Xu YH, Yue L, Wei S, Liu L, Gan X, Zhou S, Nie S (2010). Evaluating the risk of hypertension using an artificial neural network method in rural residents over the age of 35 years in a Chinese area. *Hypertens Res*, 33(7): 722-6.
9. Hsieh MH, Sun LM, Lin CL, Hsieh MJ, Hsu CY, Kao CH (2018). Development of a prediction model for pancreatic cancer in patients with type 2 diabetes using logistic regression and artificial neural network models. *Cancer Manag Res*, 10: 6317-24.
10. Rasmy L, Wu Y, Wang N, et al (2018). A study of generalizability of recurrent neural network-based predictive models for heart failure onset risk using a large and heterogeneous EHR data set. *J Biomed Inform*, 84: 11-16.