JKMS

Original Article
Obstetrics & Gynecology

Check for updates

Machine Learning Approaches to Identify Factors Associated with Women's Vasomotor Symptoms Using General Hospital Data

Ki-Jin Ryu [ID],[1] Kyong Wook Yi [ID],[1] Yong Jin Kim [ID],[1] Jung Ho Shin [ID],[1] Jun Young Hur [ID],[1] Tak Kim [ID],[1] Jong Bae Seo [ID],[2,3] Kwang-Sig Lee [ID],[4] and Hyuntae Park [ID] [1]

[1]Department of Obstetrics & Gynecology, Korea University College of Medicine, Seoul, Korea
[2]Department of Biosciences, Mokpo National University, Muan, Korea
[3]Department of Biomedicine, Health & Life Convergence Science, Mokpo National University, Muan, Korea
[4]AI Center, Korea University College of Medicine, Seoul, Korea

ORCID iDs
Ki-Jin Ryu [ID]
https://orcid.org/0000-0002-3013-8284
Kyong Wook Yi [ID]
https://orcid.org/0000-0001-7059-640X
Yong Jin Kim [ID]
https://orcid.org/0000-0001-8684-2126
Jung Ho Shin [ID]
https://orcid.org/0000-0001-7559-5457
Jun Young Hur [ID]
https://orcid.org/0000-0001-5480-7203

ABSTRACT

Background: To analyze the factors associated with women's vasomotor symptoms (VMS) using machine learning.

Methods: Data on 3,298 women, aged 40–80 years, who attended their general health check-up from January 2010 to December 2012 were obtained from Korea University Anam Hospital in Seoul, Korea. Five machine learning methods were applied and compared for the prediction of VMS, measured by the Menopause Rating Scale. Variable importance, the effect of a variable on model performance, was used for identifying the major factors associated with VMS.

Results: In terms of the mean squared error, the random forest (0.9326) was much better than linear regression (12.4856) and artificial neural networks with one, two, and three hidden layers (1.5576, 1.5184, and 1.5833, respectively). Based on the variable importance from the random forest, the most important factors associated with VMS were age, menopause age, thyroid-stimulating hormone, and monocyte, triglyceride, gamma glutamyl transferase, blood urea nitrogen, cancer antigen 19-9, C-reactive protein, and low-density lipoprotein cholesterol levels. Indeed, the following variables were ranked within the top 20 in terms of variable importance: cancer antigen 125, total cholesterol, insulin, free thyroxine, forced vital capacity, alanine aminotransferase, forced expired volume in 1 second, height, homeostatic model assessment for insulin resistance, and carcinoembryonic antigen.

Conclusion: Machine learning provides an invaluable decision support system for the prediction of VMS. For managing VMS, comprehensive consideration is needed regarding thyroid function, lipid profile, liver function, inflammation markers, insulin resistance, monocyte count, cancer antigens, and lung function.

Keywords: Vasomotor Symptoms; Hot Flashes; Menopause Age; Thyroid Stimulating Hormone; Monocyte; Cancer Antigen

INTRODUCTION

Vasomotor symptoms (VMS), referring to hot flashes and sweating, are major symptoms of peri-menopausal and post-menopausal women and a main cause of their hospital visit.[1] Once