**Presentation title:** Predicting Metastasis in Gastric Cancer Patients using Machine Learning-based Approaches

**Corresponding Author name:** Nasrin Borumandnia

**Affiliation:** Urology and Nephrology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

**Ph. No:** +989151220632

**Email ID’s:** borumand.n@gmail.com, nasrin.borumand@sbmu.ac.ir

**WhatsApp No:** +989151220632

**Any alternative number:** +989120711140

**Other Authors if any:** Atefeh Talebi1, Carlos A Celis-Morales1,2, Nasrin Borumandnia3\*, Mohamad Amin Pourhoseingholi4, Abolfazl Akbari5

1. British Heart foundation Cardiovascular Research Centre, University of Glasgow, Glasgow, UK; Email: Atefeh.Talebi@glasgow.ac.uk

2. Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK, Email: carlos.celis@glasgow.ac.uk

3\*. Urology and Nephrology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: borumand.n@gmail.com, nasrin.borumand@sbmu.ac.ir

4. Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: amin\_phg@yahoo.com

5. Colorectal Research Center, Iran University of Medical Sciences, Tehran, Iran, Email:akbariia2006@gmail.com

**Presentation type:** (Oral presentation/ Poster presentation)

**Abstract:**

Gastric cancer (GC), with a 5‑year survival rate of less than 40%, is known as the fourth leading cause of cancer-related mortality worldwide. The aim of this study is to develop predictive models based on both demographic and clinical variables using different machine learning (ML) classifiers to predict the metastasis status of patients with GC.

The data used in this study consisted of 733 GC patients, divided into train and test groups at a ratio of 8:2, who were diagnosed at Taleghani tertiary hospital. To predict metastasis in GC, ML-based algorithms, including Naive Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), Neural Network (NN), Decision Tree (DT) and Logistic Regression (LR), were performed with 5-fold cross-validation. The model performance was assessed by F1 score, precision, sensitivity, specificity, area under the curve (AUC) of receiver operating characteristic (ROC) curve and precision-recall AUC (PR AUC).

Among 733 patients with GC, 262 (36%) experienced metastases. Although all models had optimal performance, the indices of SVM model seemed to be more appropriate (training set: AUC: 0.94, Sensitivity: 0.94; testing set: AUC: 0.85, Sensitivity: 0.92). Then, Next, NN had the highest AUC among ML approaches (training set: AUC: 0.98; testing set: AUC: 0.86). The RF model, which determined tumor size and age as two essential variables, is considered as the third efficient model, due to higher specificity and AUC (84% and 87%).

Based on the demographic and clinical characteristics, ML approaches can predict the metastasis status of GC patients. According to AUC, sensitivity and specificity, both SVM and NN can be regarded as better algorithms among the six applied ML-based methods.

**Biography:**

I am Dr. Nasrin Borumandnia, an assistant professor of biostatistics at Shahid Beheshti University of Medical Sciences. I have a PhD in biostatistics from the same university, and I have been working as an assistant professor since 2017, teaching courses on statistical methods and data analysis. I have designed and analyzed data from various types of studies, including clinical trials, observational studies, surveys, fMRI data, and big data, using a range of statistical techniques, from basic to advanced methods such as multivariate analysis and causal inference. I am also proficient in machine learning methods, such as random forest, neural network, and support vector machine, which I have applied to perform data mining, classification, and prediction tasks. I have published more than 70 papers in peer-reviewed journals.