

## Predicting the Incidence and Trend of Breast Cancer Using Time Series Analysis for 2007-2016 in Qazvin

Hajiabadi F<sup>1</sup>, Bagheri H<sup>1</sup>, Tonokaboni N<sup>2</sup>, Zamanian M<sup>3</sup>, Hosseinkhani Z<sup>4\*</sup>

<sup>1</sup> Children Growth Research Center, Research Institute for Prevention of Non-Communicable Diseases, Qazvin University of Medical Sciences, Qazvin, Iran

<sup>2</sup> Non-Communicable Diseases Control Office, Deputy of Health, Qazvin University of Medical Sciences, Qazvin, Iran

<sup>3</sup> Department of Epidemiology and Biostatistics, School of Health, Arak University of Medical Sciences, Arak, Iran

<sup>4</sup> Metabolic Diseases Research Center, Research Institute for Prevention of Non-Communicable Diseases, Qazvin University of Medical Sciences, Qazvin, Iran

Receive: 19/8/2020

Accepted: 4/2/2021

\*Corresponding Author:  
zhosseinkhani122@gmail.com

Ethics Approval:  
Not Applicable

### Abstract

**Introduction:** Breast cancer is the most common cancer and the second leading cause of death in women worldwide. The aim of this study was to analyze the trend and predict the incidence of breast cancer using time series analysis.

**Methods:** In this study, data on breast cancer incidence in Qazvin province between 2007 and 2016 were analyzed using time series analysis with autoregressive integrated moving average (ARIMA) modeling to forecast the future pattern. The Box-Jenkins time series model and its diagnosis and evaluation methods were used to show the trend and forecasting the next year new cancers. To describe and fit the appropriate models, R statistical software version 3.6.3 was used.

**Results:** Between 2007 and 2016, a total number of 1229 new patients had been registered (monthly mean [SD]: 10.24 [1.03]). Although the overall trend in the raw number of new breast cancer cases has been increasing over time, the change in observations over time has been increasing and decreasing. According to Bartlett test results, the variances of the data were not constant. Also, according to the results of Kolmogorov-Smirnov test, breast cancer series data were not normal. Among the studied models, ARIMA (1, 1, 1) was selected due to lower AIC criteria than other models, and this model was selected as the final model for predicting breast cancer for the next year. The confidence interval of the predicted values was relatively narrow, which indicates the appropriateness of the final model in the prediction.

**Conclusion:** Time series analysis is an efficient tool to model the past and future data on the raw number of new cancer cases, and the goodness-of-fit indicators of the model showed that the Box-Jenkins model is a reliable model for fitting similar data.

**Keywords:** Breast Cancer, Seasonal Trend, Time Series Analysis, Iran