ORIGINAL ARTICLE

**Iranian Quarterly Journal of Breast Disease 2016; 9(1).**

# Comparison and Evaluation of Synthesis of Risk Factors in Breast Cancer and Provide a Model for Determine the Likelihood of Developing Breast Cancer Using by EM Algorithm in Data Mining Techniques

**Hemmasian Etefagh M:** Faculty of Computer Engineering, Najafabad branch, Islamic Azad University, Najafabad, Isfahan, Iran
**Nadimi Shahraki MH:** Faculty of Computer Engineering, Najafabad branch, Islamic Azad University, Najafabad, Isfahan, Iran

**Corresponding Author:** Mahdi Hemmasian Ettefagh, Mahdi_hemmasian@yahoo.com

[ Downloaded from ijbd.ir on 2026-11-01 ]

[ DOR: 20.1001.1.17359406.1395.9.1.3.5 ]

## Abstract

**Introduction:** Anything that increases a person's chance of developing breast cancer is risk factor. Awareness of these risk factors to identify women who have a high risk of developing breast cancer may be helpful. It also allows intervening in individual and social risk factors and adjustment the incidence of cancer in the individual and our community. The aim of this study is to provide a mathematical model to determine the risk of breast cancer in screening centers.

**Methods:** In this study we used the data from Breast Cancer Surveillance Consortium (BCSC) in National Cancer Institute in USA from 2000-2009 on 6318638 people. For clustering we used EM algorithm and executed in Weka software to cluster this data.

**Results:** All of the risk factors do not increase the risk of breast cancer in a same mode. Some of these factors are more dangerous and increases the risk raise more than from another risk. Therefore, to determine the impact factor of the population could be achieved in a model upon which to predict the likelihood of breast cancer will be there.

**Conclusion:** The results of this study shown that using EM algorithm and effective clustering may be effective, any of the risk factors studied in a dataset and for each factor as effective coefficient obtained, and finally the sum of these impact factors on the risk of breast cancer risk factors, the risk can be predicted.

**Keywords:** Breast Cancer, Risk Factor, Data Mining, Screening, EM Algorithm.