COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 122 (2015) 123–135

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A risk management model for familial breast cancer: A new application using Fuzzy Cognitive Map method

Elpiniki I. Papageorgiou [a], *Jayashree Subramanian* [b], *Akila Karmegam* [c], *Nikolaos Papandrianos* [d,*]

[a] *Computer Engineering Department, Technological Educational Institute of Central Greece, 3rd KM Old National Road Lamia-Athens, 35100 Lamia, Greece*
[b] *Computer Science Engineering, RVS College of Engineering and Technology, Coimbatore, India*
[c] *Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, India*
[d] *Diagnostic Medical Center "Diagnostiki A.E.", Larisa, Greece*

ARTICLE INFO

Article history:
Received 2 March 2015
Received in revised form 1 July 2015
Accepted 3 July 2015

Keywords:
Familial breast cancer
Risk assessment
Fuzzy Cognitive Maps
Hebbian learning
Medical decision support systems

ABSTRACT

Breast cancer is the most deadly disease affecting women and thus it is natural for women aged 40–49 years (who have a family history of breast cancer or other related cancers) to assess their personal risk for developing familial breast cancer (FBC). Besides, as each individual woman possesses different levels of risk of developing breast cancer depending on their family history, genetic predispositions and personal medical history, individualized care setting mechanism needs to be identified so that appropriate risk assessment, counseling, screening, and prevention options can be determined by the health care professionals. The presented work aims at developing a soft computing based medical decision support system using Fuzzy Cognitive Map (FCM) that assists health care professionals in deciding the individualized care setting mechanisms based on the FBC risk level of the given women. The FCM based FBC risk management system uses NHL to learn causal weights from 40 patient records and achieves a 95% diagnostic accuracy. The results obtained from the proposed model are in concurrence with the comprehensive risk evaluation tool based on Tyrer–Cuzick model for 38/40 patient cases (95%). Besides, the proposed model identifies high risk women by calculating higher accuracy of prediction than the standard Gail and NSAPB models. The testing accuracy of the proposed model using 10-fold cross validation technique outperforms other standard machine learning based inference engines as well as previous FCM-based risk prediction methods for BC.

* *Corresponding author*. Tel.: +30 6936064613.
E-mail addresses: epapageorgiou@teiste.gr, e.i.papageorgiou75@gmail.com (E.I. Papageorgiou), jayashreecls@gmail.com (Jayashree Subramanian), skylabm@gmail.com (A. Karmegam), nikpapandr@gmail.com (N. Papandrianos).

http://dx.doi.org/10.1016/j.cmpb.2015.07.003
0169-2607/

1. Introduction

Breast Cancer (BC) is the second most common malignancy among women with an estimated 1.38 million new cases per year globally [1]. It develops in the breast tissue, primarily in the milk ducts (ductal carcinoma) or glands (lobular carcinoma) which begin with the formation of a small tumor or mass [2]. A mass with a smooth, well defined border is non cancerous (benign). Mass with an irregular border or with spiculations may be cancerous (malignant). Detection and diagnosis of breast cancer in an early stage increases the chances for successful treatment and complete recovery of the patient.

Familial breast cancer typically occurs in people with an unusually high number of family members affected by breast, ovarian or a related cancer [3]. Women having a number of first and/or second degree relatives affected with the above mentioned types of cancers either in their maternal or paternal lineage, have a high probability of inheriting the BRCA1/BRCA2 gene mutations. These people have an increased future risk of developing malignant breast tumors; thus they need to be given special attention in terms of the screening procedures, the screening frequency and other follow up procedures, even though their current screening study is absolutely normal and they remain asymptomatic. Hence, there is a strong need for developing a familial breast cancer risk assessment model that takes into account the risk factors related to age, personal medical history of the individual and more importantly her family history of breast cancer related incidences, as well as the demographic factors. This is expected to give a formal FBC risk assessment.

National Cancer Institute has developed an interactive risk assessment tool, namely Breast Cancer Risk Assessment Tool (BCRAT) which is based on the GAIL model [4]. The Gail model considers only first degree relatives and also takes no account of the age of onset of breast cancer. It underestimates the risk in the 50% of families with cancer in the paternal lineage [5]. As compared with the Gail model, the Claus model incorporates more extensive information about family history, but it excludes risk factors except family history [6,7]. Similar to the Gail model, the Claus model does not also consider the age and life style related factors for estimating the risk. Challa et al. (2013) have extended the Gail Model by including the puberty age and number of biopsies taken [8] to the Indian population in order to assess the probability of breast cancer. As it refers to the BRCAPRO model, it calculates the probability occurrence of a germ-line mutation of the BRCA1 and BRCA2 genes [9]. Moreover, the Tyrer and Cuzick model [10] integrates the factors related to family history, hormonal factors, and benign breast disease. These different models give aid to different categories of individuals by estimating the risk of developing a malignant tumor and provide suggestions for the required life style changes which could decrease their risk [10–13].

Therefore, the available BC risk assessment tools/models concentrate either on family history, genetic mutations or hormonal factors and after all, these models perform well for populations but poorly for individuals. The models that are designed based on personal history (Gail model and its extensions) do not cover sufficiently the family history. Other models like Claus that focus on family history, perform poorly when personal history is added. Models like BRCAPRO only take into account the genetic mutations. There are only a few models like Tyrer–Cuzick that are comprehensive enough for BC risk prediction.

In this article, a soft computing methodology is used to develop a familial BC risk assessment model which concentrates on both family history and personal medical history including also the demographic factors described in literature. FCM belongs to the class of soft computing techniques that is useful to model the dynamics involved in a given complex systems using a set of concepts and the relationship with each other. FCM inherits the main aspects of fuzzy logic and neural networks in a graph-based structure modeling complex decision making problems. Due to their ease to construct and use, their flexibility and adaptation to practically any problem domain, their support on uncertain knowledge, their relatively simple and comprehensible modeling philosophy which is very close to human reasoning and their capabilities to handle efficiently complex issues in different domains, FCM has become popular and found large applicability to many diverse scientific areas from knowledge modeling to prediction and decision making [14–17,55–57].

In a recent previous work for medical decision support [18], an integrated risk prediction model using a two-level Fuzzy Cognitive Map (FCM) structure was developed. The proposed model combined the results of the initial screening mammogram of the given woman with her demographic risk factors to predict the post-screening risk of developing BC. More specifically, the Level-1 FCM models the demographic risk profile and through Hebbian learning it helps on predicting the BC risk grade based on demographic risk factors. The risk grades estimated by the proposed model were validated using two standard BC risk assessment models viz. Gail and Tyrer–Cuzick. The Level-2 FCM models the features of the screening mammogram concerning normal, benign and malignant cases. The data driven hebbian learning algorithm (DDNHL) was used to train this model in order to predict the BC risk grade based on these mammographic image features. An overall risk grade was calculated by combining the outcomes of these two FCMs.

Even though the proposed L1-FCM model included the demographic risk factors identified by domain experts, it does not completely cover the history of the individual, mainly the family history and personal medical history. The purpose of this study is to develop a BC risk assessment model which concentrates on family history and personal medical history using FCM methodology enhanced by its Hebbian learning capabilities. Moreover it aims to compare it with the standard models which estimate the risk of familial breast cancer among women. Actually, in this research study, we further consider more features related to the family history, such as the family history of first and second degree relatives, male breast cancer and personal medical history of the individual person in deciding their FBC risk level. The proposed model is concentrated on individualized personal decision making and though the efficient capabilities of FCMs can provide effective decisions about the individual BC risk level. No previous research has explored FCM for the stated purpose.