ORIGINAL PAPER

Breast Alert: An On-line Tool for Predicting the Lifetime Risk of Women Breast Cancer

Joel J. P. C. Rodrigues · Nuno Reis ·
José A. F. Moutinho · Isabel de la Torre

Abstract Breast Cancer is an important disease that affects many women, excluding self-examination and screening by mammogram, nothing helps women or their physicians to know what risk they run of suffering from breast cancer during the course of their lives. There have been many studies detailing the relative risks of breast cancer based on different factors and applications to calculate the breast cancer risk, but none implemented in a way to show lifetime risk. This paper presents an on-line tool (called Breast Alert) to calculate the lifetime breast cancer risk for women using a proposed model. With Breast Alert, physicians can make a quick screening for women when they consult. It is easy to use and intuitive. In a few minutes, physicians can have a lifetime breast cancer risk. This tool does not replace tests like self-examination, breast screening or detection by other options, but allows for the proper precautions to be taken and calls attention to the expected lifetime risk. Nowadays, 300 women (between 20 and 75 years old) from different countries have used the system and most of them (80%) have a higher than normal chance of contracting breast cancer. With these results, it is important to alert of the importance to make an early prevention of breast cancer in different women groups.

Keywords Breast cancer · e-Health · Java · Lifetime risk · Web technologies

J. J. P. C. Rodrigues (✉) · N. Reis
Instituto de Telecomunicações, University of Beira Interior,
Rua Marquês d'Ávila e Bolama,
6201-001 Covilhã, Portugal
e-mail: joeljr@ieee.org

N. Reis
e-mail: nuno.reis@it.ubi.pt

J. A. F. Moutinho
Health Sciences Faculty, University of Beira Interior,
Rua Marquês d'Ávila e Bolama,
6201-001 Covilhã, Portugal
e-mail: jafmoutinho@fcsaude.ubi.pt

I. de la Torre
Department of Signal Theory and Communications,
University of Valladolid,
Campus Miguel Delibes, s/n,
47011 Valladolid, Spain
e-mail: isator@tel.uva.es

Introduction

Breast cancer is the most common type of cancer in women and the second leading cause of death by cancer. Although scientists do not know the exact causes of most breast cancer, they do know some of the risk factors that increase the likelihood of a woman developing breast cancer in her life [1–4]. Some of them are genetic and relate to family history, others are based on personal factors such as reproductive history and medical record [5]. Breast Cancer is a disease affecting many women around the world, excluding self-examination and screening by mammogram, nothings helps women or their physicians to know what risk they run of suffering from breast cancer during the course of their lives.

Several research studies have analyzed sets of breast cancer data related to breast cancer diagnosis and addressed the prediction of breast cancer outcomes, further research into this field will enable patients to have an idea of the prognosis of the likely course and outcome of their disease [6–8].

There have been many studies detailing the relative risks of breast cancer based on these factors. The most widely used breast cancer risk prediction tool is the Gail et al. model [9]. Its parameterization and a thorough history of its development are discussed elsewhere [10]. Tyrer et al. (2004) developed a model incorporating the Breast Cancer

(BRCA) genes, a low penetration gene and personal risk factors. This model is one of the most up-to-date and offers a high degree of reliability [11].

Moreover, there are several on-line tools to predict the breast cancer in women but none implemented in a way to show lifetime risk. This paper presents an on-line tool (Breast Alert) to calculate the lifetime risk of breast cancer using the proposed model in [11]. With this system, and information such as current age, age at first menstruation, age of first born child, height, weight and other personal and family factors, based on some published studies [11], where we base this work can reach percentages of risk over the years that will help a lot for early breast cancer detection. Physicians can use this tool to make a quick screening of all women they consult and they can have a lifetime risk for women. It is important to underline that this tool does not intend to replace with this system other tests like self-examination, breast screening or detection by other options, but to alert the woman to take the proper precautions, and calling attention to the expected risk. With this system we want to help women and physician to calculate her lifetime risk and take the necessary precautions, because an early detection can save a life.

The remainder of this paper is organized as follows. Section Related Work elaborates on related literature about available webpages to predict breast cancer risk. Section Methodology describes the followed methodology to create the proposed platform while Section Results focuses on the results achieved and describes its main modules (the administrator module, the calculations and user statistics). Finally, Section Discussion and Conclusions discusses results and concludes the paper.

Related work

This section presents some of the research literature related to websites to predict breast cancer risk. Following are some of these websites:

– Siteman Cancer Center [12]. This test was developed by the Harvard Center for Cancer Prevention. The results are given as "low", "medium", or "high" risk. It includes tips to lower your risk, positive feedback on things that you currently do which lower your risk and the users have many other available e-health tests. Also, the users do not have a lifetime risk, the test is hard to complete, the Web page is not attractive and the test must be done in English.
– Dr. Steven Halls Detailed Breast Cancer Risk Calculator [13]. This test was created by Steven B. Halls, MD. It has 6 questions, with 6 additional modifier questions. It has clear and simple language, with explanations. It

uses the National Surgical Adjuvant Breast and Bowel Project (NSABP) Model. The results are given the risk as a percentage. It is easy to answer the test. The website has a note saying that other risk factors exist, but they are not included in the calculator. This is a poor approach; if they don't consider all risk factors the veracity of the results could be compromised. The way they show results is very poor. They only show a number for risk in the next 5, 10, 20 and 30 years. Most of people get nothing from this test, because they cannot understand the meaning of the numbers. It does not have a lifetime risk graphic and the test must be done in English.

– National Cancer Institute Breast Cancer Risk Assessment Tool [14]. The Breast Cancer Risk Assessment Tool is an interactive tool designed by scientists at the National Cancer Institute (NCI) and the NSABP. It has 7 questions with links to glossary terms. The results are given in as a percentage chance. It gives a comparison of your results to women in the U.S. who are in your same age and ethnic group. It was designed for use by health professionals. It is based on the Gail Model. It is easy to answer the test. Other risk factors exist, but they are regrettably not included in this calculation. It is useful only for women of 35 years or over. It does not have a lifetime risk graphic and the test must be done in English. Some risk factors are not considered in the test. Once again, the way the information is given is poor. They show the woman's in risk in 5 years time and at 90 years old and compare this with the average.
– Breast Cancer Prevention's Risk Assessment [15]. This Website is a product of the NSABP. It is a complete test. The results are given as a percentage chance and are validated for women from 35 to 84 years old. It does not have a lifetime risk graphic and the test must be done in English. The results are wordy, but worth reading, with some useless questions. The way the information is given is poor. They show the woman's risks in 5 years time and at 80 years old and compare this with the average.

These websites are not easy to use and the result does not allow one to determine with ease the real risk or the lifetime risk.

Methodology

This section presents the used breast cancer prediction model, the system overview, and its architecture.

Breast cancer prediction model

There are many factors for determining a woman's risk of breast cancer. Some of them are genetic and related to

family history; others are based on personal factors. In this work, we employed the proposed model in [11]. That approach was taken to use a two locus genetic model with one of the genes based on BRCA1 and BRCA2. The authors used Bayes theorem to calculate the genetic probability based on the family history and estimated the population frequency of these genes. Inherited mutations in the BRCA1 and BRCA2 genes confer significantly increased lifetime risks of breast cancer [16]. A meta-analysis of 22 studies of BRCA1 and BRCA2 mutation carriers who were unselected for family history was carried out [17]. The average cumulative risk of Breast cancer up to 70 years of age was 65% in BRCA1 carriers and 45% in BRCA2 carriers. Moreover, BRCA1 and BRCA2 mutations are found in approximately 5% of all breast cancers and in up to 20–25% in the case of a family history of breast and/or ovarian cancer [17–19].

According to [11], for a woman the risk of developing breast cancer between ages t_1 and t_2 is given by Eq. 1.

$$Probability\ (cancer) = 1 - \left(1 - \sum_{i=1}^{6} p_i F_i(t_1, t_2)\right)^{\alpha} \quad (1)$$

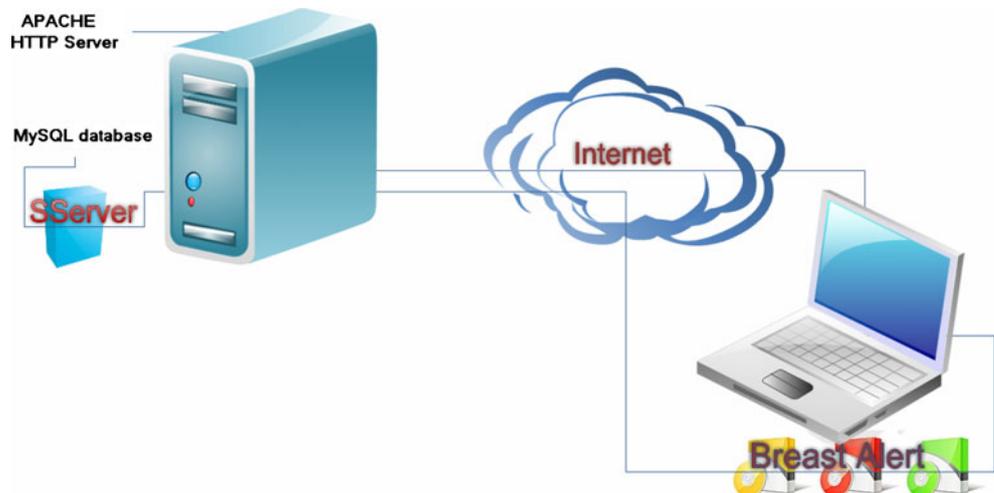Where,

p_i is the probability of the woman having the relevant phenotype. The woman's family history is employed to calculate the distribution of her genotype probabilities and from this the phenotypic probabilities are calculated.

$F_i (t_1; t_2)$ is the probability of getting breast cancer between ages t_1 and t_2 given the woman's phenotype i.

α is the relative risk due to personal factors.

System overview and architecture

The Breast Alert system is multi-language, so it will be easy to use for any woman. It asks for some personal and family factors that will be used as a basis for the calculation.

After obtaining information about personal and family history, based on the studies shown in [11], a method is arrived for calculating the lifetime risk of breast cancer. There is a database where all information will be saved and will be available to those with permission, whenever they want.

In Breast Alert application, a standard calculation is also carried out with the probabilities of having breast cancer, which is based on the study [11], but there is also the possibility of the physician adjusting the probabilities data and saving it with another name. Figure 1 shows the application architecture using three-layered approach. A free open-source application and database servers were chosen and a relational database to store the personal data was selected.

The chosen technologies and the reason for making these choices are described in the following lines. For the Web Server, Httpwebserver–Apache/2.2.14 (Apache) was chosen. The web server will run a very simple page, made in Hypertext Transfer Protocol (HTTP) and PHP: Hypertext Preprocessor (PHP).

The Database Management System (DBMS) used was MyStructured Query Language (MySQL) because it is one of the World's most popular open source DBMS, and provides consistent fast performance, high reliability, and ease to use.

The choice of Java was due to be an object oriented language and independent platform. The initial objective was to provide access to the application via Website; many other technologies could be used perfectly.
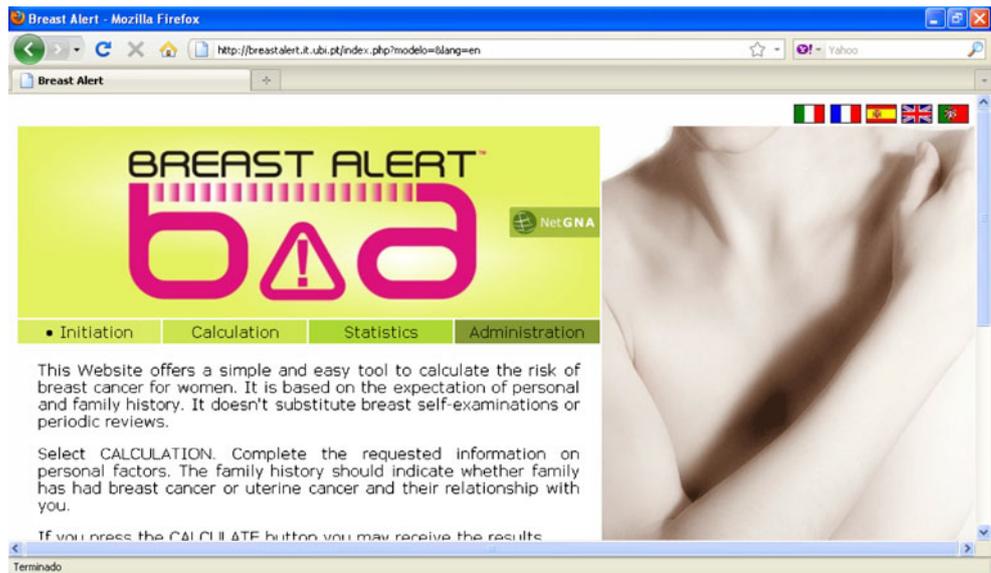
Results

In this section, the following is laid out: the manager module, the calculation and the system evaluation.

Manager module

Administrator program is used to create and change the models where the calculations were made. The system is

Fig. 1 Breast Alert system architecture

available through http://breastalert.it.ubi.pt. In Fig. 2 can
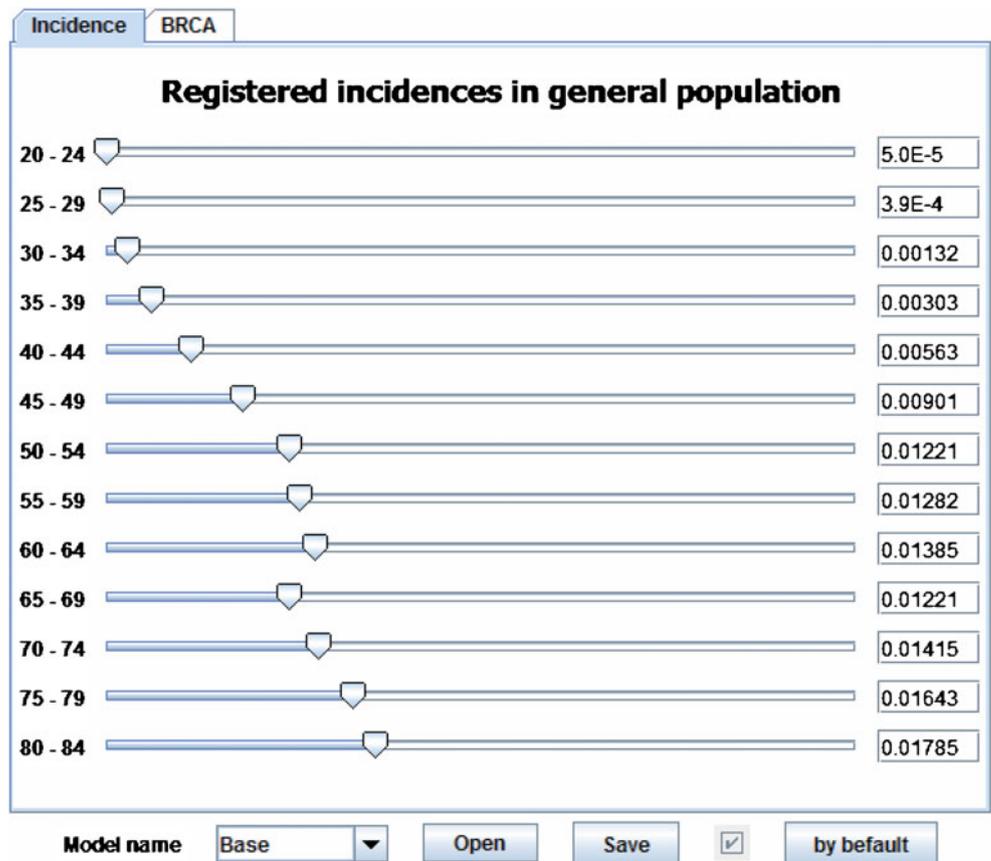be seen the multilingual presentation platform of the
system.

This is a fundamental part of the programme, in this
interface user name and password are required to enter. In this
interface, we will show the probability of having breast cancer

in 5 years intervals from 20 to 85 years old. The actual data
came from the incidence of breast cancer in the UK population
[20]. This data can be changed any time with any others
published statistics data. In this interface is also possible to
change the BRCA1 and BRCA2, the incidence table of these
genes having been obtained from the same study [11] and any

Fig. 3 Manager system

Fig. 4 Form to insert personal issues

change in these values should carry the same precautions as the incidence table. This interface allows many functions, they are:

– Create new models.
– Change values in the models.
– Choose the default model for calculation.

These values are the base for all calculations. Any change in these values can change drastically the results of the lifetime risk calculation. The interface was designed for its ease of use; it can be done by entering the value in the box or just sliding the bar. After all changes have been made, they can be saved in the current model or in a new
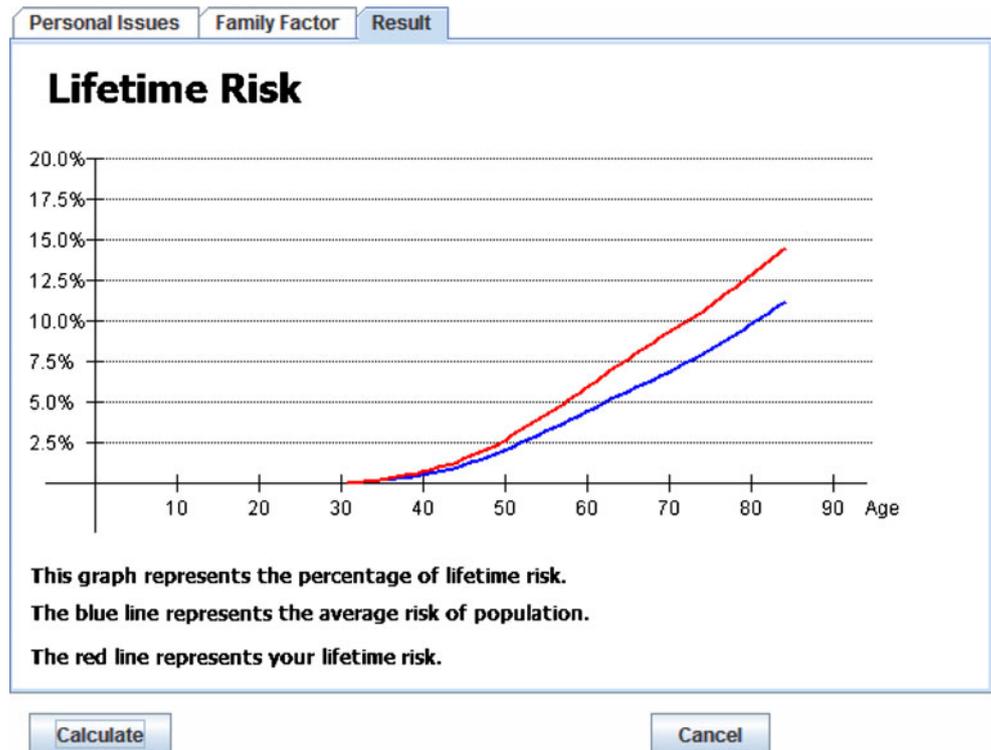
Fig. 5 Form to add family factors

Fig. 6 Example of a test result showing the lifetime risk

model. In Fig. 3 can be seen a snapshot of the system management, with the aim of analyzing those women who have taken the test and the results obtained concerning the relative risk of contracting breast cancer during the course of their lives.

Breast alert calculator

In order to calculate the risk of breast cancer during the course of a life, the women must introduce personal and family information into the Breast Alert platform. In Fig. 4

Fig. 7 Database with people who have carried out the test

Date	e-mail	Age	Risk in 10 years ▼	Risk in 80 years
23-06-2010		35	3,342	28,249
22-06-2010		65	2,715	4,244
24-06-2010		52	2,506	7,155
28-06-2010		46	2,417	9,625
22-06-2010		60	2,373	4,546
02-07-2010		47	2,352	8,825
24-06-2010		38	1,921	15,029
01-07-2010		41	1,539	9,135
21-06-2010		36	1,485	14,387
22-06-2010		39	1,437	10,316
24-06-2010		38	1,393	10,947
18-06-2010		35	1,324	14,488
22-06-2010		40	1,292	8,53
22-06-2010		36	1,076	10,357
22-06-2010		35	0,996	10,803
28-06-2010		36	0,918	8,779
29-06-2010		35	0,846	9,077
28-06-2010		34	0,829	9,897
24-06-2010		32	0,811	12,492
23-06-2010		35	0,81	8,667
23-06-2010		33	0,808	10,865
23-06-2010		33	0,743	9,925
23-06-2010		31	0,697	12,413

can be seen the interface where the women introduce their personal information. This interface is friendly for the user. The information relating to the family medical history of ovarian and/or breast cancer is introduced in the interface which can be seen in Fig. 5. For example, one can see here that the mother suffered from breast cancer (red circle) and the maternal grandmother from ovarian cancer (blue circle).

The results obtained from this test can be consulted in the "Result" tag of the system. Figure 6 includes a graphical representation of the breast cancer risk during the course of her life. An e-mail is sent to users when they complete the test, if her email address is inserted on the personal issues form.

The ages of the patient appear in the x axis and the percentage chance of risk of contracting breast cancer during the course of her life appears in the y axis (the blue line represents the average for the population and the red line is that of the patient). For example, in this case we can see that the patient has a higher than average risk of suffering from breast cancer during the course of her life. Therefore, with this information, she can take the necessary precautions and methods to changing lifestyle or eating habits, avoiding things known to cause cancer, etc.

User statistics

In this section we show the results obtained up to today from those people registered in the system. Three hundred women, between 20 and 75 years of age, used the system. In Fig. 7 can be seen a group of women who have used the on-line tool. The average of these women is 42 years and the majority (80%) have a higher than average risk of suffering from breast cancer. With the passing of the years it is assumed that this risk increases. In the majority of these cases the risk at 80 years of age grows exponentially. With these results, it is important to alert to women of an early prevention of breast cancer. To prevent new cancers from starting, scientists look at risk factors and protective factors. Some risk factors for cancer can be avoided, but many cannot. Regular exercise and a healthy diet may be protective factors for some types of breast cancer.

Discussion and conclusions

This paper presents an on-line application, Breast Alert, to allow early breast cancer prediction and route women with high risk to an appointment with a specialist physician and this was totally achieved. The Website (http://breastalert.it.ubi.pt) is available in different languages (Portuguese, English, French, Spanish and Italian). The results are given in lifetime graphic risk. They are compared with a standard lifetime graphic risk. It is easy to compare the personal lifetime risk with the standard lifetime risk.

The Breast Alert programme does not exclude self-examination or screening by mammogram, but it may help both the women and their physicians to make an early diagnosis breast cancer with help of the lifetime risk calculation in contrast to other tools.

In the future, a mobile version of Breast Alert will be created, so the physicians can easily carry in their mobile a Breast Alert version to use anywhere. Another improvement is to compare and implement other Breast Cancer assessment models. In this way, it would be possible to create multiple lifetime risks. This would create a greater variety of assessment models and with the help of the physician following the woman's history, they can decide which of the models is better.

With the physician's help, if we can follow some patients, new parameters can be discovered to improve the lifetime breast cancer risk.

References

1. Castro, K. J., Kessler, E., Erickson, B. A., and Kseibi, S. A., The relation of breast cancer staging to screening protocol compliance: A computer simulation study. *Comput. Biol. Med.* 35:91–101, 2005.
2. Carter, K. J., Kessler, E., Erickson, B. A., and Kseibi, S. A., A computer model for the study of breast cancer. *Comput. Biol. Med.* 33:345–360, 2003.
3. Chen, T.-C., and Hsu, T.-C., A GAs based approach for mining breast cancer pattern. *Expert Syst. Appl.* 30(4):674–681, 2006.
4. Delen, G. W., and Kadam, A., Predicting breast cancer survivability: A comparison of three data mining methods. *Artif. Intell. Med.* 34:113–127, 2005.
5. Amir, E., Evans, D. G., Shenton, A., Lalloo, F., Moran, A., Boggis, C., Wilson, M., and Howell, A., Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. *J. Med. Genet.* 40(11):807–814, 2003.
6. Thongkam, J., Xu, G., Zhang, Y., and Huang, F., Toward breast cancer survivability prediction models through improving training space. *Expert Syst. Appl.* 36:12200–12209, 2009.
7. Thongkam, J., Xu, G., Zhang, Y., and Huang, F., Support vector machines for outlier detection in cancers survivability prediction. International workshop on health data management. APWeb'08, pp. 99–109, 2008.
8. Thongkam, J., Xu, G., and Zhang, Y., An analysis of data selection methods on classifiers accuracy measures. *Journal of Korn Ken University* 35:1–10, 2008.
9. Pharoah, P., Gail et al risk prediction model adequately estimates breast cancer risk in groups of women but has lower accuracy for individual women. *Evid.-based Oncol.* 3(1):12–13, 2002.
10. Rockhilla, B., Byrneb, C., Rosnerc, B., Louiec, M. M., and Colditzc, G., Breast cancer risk prediction with a log-incidence model: Evaluation of accuracy. *J. Clin. Epidemiol.* 56:856–861, 2003.

11. Tyrer, S. W., and Cuzick, J., A breast cancer prediction model incorporating familial and personal risk factors. *Stat. Med.* 23:1111–1130, 2004.
12. Siteman Cancer Center [Online]. Available: http://www.siteman.wustl.edu.
13. Dr Steven Halls Detailed Breast Cancer Risk Calculator. [Online]. Available: http://www.halls.md/breast/risk.htm.
14. National Cancer Institute Breast Cancer Risk Assessment Tool. [Online]. Available: http://www.cancer.gov/bcrisktool.
15. Breast Cancer Prevention's Risk Assessment. [Online]. Available: http://www.breastcancerprevention.org.
16. Dennis, J., Ghadirian, P., Little, J., Lubinski, J., Gronwald, J., Kim-Sing, C., Foulkes, W., Moller, P., Lynch, H., Neuhasen, S., Domchek, S., Armel, S., Isaacs, C., Tungm, N., Sweet, K., Ainsworth, P., Sun, P., Krewski, D., Norad, S., Alcohol consumption and the risk of breast cancer among BRCA1 and BRCA2 mutation carriers. The Breast, pp. 1–5, 2010.
17. Antoniou, A., Pharoah, P. D., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J. L., et al., Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: A combined analysis of 22 studies. *Am. J. Hum. Genet.* 72:1117–1130, 2003.
18. Kirova, Y., Stoppa-Lyonnet, D., Savignoni, A., Sigal-Zafrani, B., Fabre, N., and Fourquet, A., Risk of breast cancer recurrence and contralateral breast cancer in relation to BRCA1 and BRCA2 mutation status following breast-conserving surgery and radiotherapy. *Eur. J. Cancer* 41:2304–2311, 2005.
19. Adem, C., Reynolds, C., Soderberg, C. L., et al., Pathologic characteristics of breast parenchyma in patients with hereditary breast carcinoma, including BRCA1 and BRCA2 mutation carriers. *Cancer* 97:1–11, 2003.
20. Breast cancer statistics–Key Facts. [Online]. Available: http://info.cancerresearchuk.org/cancerstats/types/breast.