Decision Support Systems for Breast Cancer *

André Fontes[1] and João Cunha[1]

Higher Institute of Engineering, Polytechnic of Porto, Porto, Portugal
1210083@isep.ipp.pt, 1171857@isep.ipp.pt

Abstract. Breast cancer is the most common type of cancer among women, and it is the second leading cause of cancer death in women. Breast cancer is one of the diseases with the greatest impact on our society, not only because it is very frequent, and associated with a very serious image, but also because it attacks an organ full of symbolism, in motherhood and femininity. In order to contribute to this broad challenge we have Decision Support Systems (DSS),Information Management (IM) and Knowledge Management (KM) techniques and tools to help professionals on a daily bases to analyze a patient correctly. The objective is to create a tool that manages and helps professionals to decision making for breast cancer.

Keywords: Breast Cancer · Decision Support Systems · Information Management.

1 Introduction

1.1 Contextualization

In Portugal, around 7,000 new cases of breast cancer are detected annually, and 1,800 women die from this disease. The signs and symptoms of cancer can variate, and some women who have cancer may not have any of these signs and symptoms. In any case, it is recommended that the woman knows her breasts, and knows how to recognize changes in order to alert the doctor. Making the right decisions towards breast cancer is very important for the patient receive the right treatment.

Clinicians face hard decisions in many aspects of breast cancer. For example, image-guided core needle biopsy of the breast is a common procedure that can be non-definitive in 5% - 15% of women. This makes it difficult to further classify breast cancer into sub-types. Variation in breast cancer sub-types has been known to be associated with a patient's drug response, progression of the tumor, and survival of the patient.[1]

Recently, the idea of "computers helping doctors" has increasingly become a reality due to the evolution of artificial intelligence (AI) technologies. The implementation of AI in CAD systems holds great promise for the future of cancer detection. These systems are designed to support first-line tumor diagnosis

* ISEP

by providing a second, potentially objective, opinion on the content of medical images.[2]

Personalized treatment strategies are critical for the improvement of the BC patients follow-up. An important step has been made in medical practice through Clinical Decision Support Systems (CDSSs). They are computerized systems that analyze several data to support clinical decision-making by improving its quality.[3]

The objective of this paper is to study problems related to decision support system´s and develop one software that helps professionals making the right decisions in based analyzes of breast cancer.

2 Body

2.1 Breast Cancer

In this chapter, you can find important information about breast cancer. Possible causes, screening, symptoms, diagnosis, treatment, and recovery will be addressed. It also contains information to help women with breast cancer cope with the disease.

Breast cancer is an uncontrolled growth of breast cells. To better understand breast cancer, it helps to understand how any cancer can develop.Cancer occurs as a result of mutations, or abnormal changes, in the genes responsible for regulating the growth of cells and keeping them healthy. The genes are in each cell's nucleus, which acts as the control room of each cell.[4]

Normally, the cells in our bodies replace themselves through an orderly process of cell growth: healthy new cells take over as old ones die out. But over time, mutations can turn on certain genes and turn off others in a cell. That changed cell gains the ability to keep dividing without control or order, producing more cells just like it and forming a tumor.[4]

2.2 Prevention

Making self exams and have routine exams is very important to act fast on the problem and to prevent breast cancer. To make sure that you don't have breast cancer should be aware of the symptoms:

- Swelling of all or part of the breast.
- Skin irritation or dimpling.
- Breast pain
- Nipple pain or the nipple turning inward.
- Redness, scaliness, or thickening of the nipple or breast skin.
- Lump in the underarm area.
- Nipple discharge other than breast milk.

Genetic can also be a factor towards having breast cancer that's why you should have consult a genetic counselor and provide the information of your background, like what relatives had cancer, what cancers they had and at what ages they were diagnosed. After the doctor have that information he can make a better assessment that someone has an increased risk of carrying a harmful BRCA1 or BRCA2 gene variant and make appropriate genetic test. People who carry that type of genetic have a higher chance to have breast cancer and to develop more harm full types of cancer.

2.3 Detection

Investigators have studied many diagnostic methods for diagnosing early-stage breast cancer, including mammography, MRI, ultrasonography, PET, breast MI and biopsy. Table 1 compares the most commonly used breast cancer screening techniques and their respective limitations.[5]

Table 1. Conventional breast screening methods and their limitations.[5]

Type	Use	Sensitivity *	Specificity *	limitations	Time
Mammography	Mass screening. Image bone, soft tissue and blood vessels all at the same time. Shadowing due to dense tissues	67.8%	75.0%	Ionizing radiation, low sensitivity and specificity, sensitivity drops with tissue density increases	few seconds
Ultrasound	Evaluate lumps found in mammography; Not suitable for bony structures	83.0%	34.0%	Low sensitivity; experienced operator is required during examination; low resolution image;	10–20 min
MRI	Young women with high risk; Images small details of soft tissues	94.4%	26.4%	Some types of cancers cannot be detected such as ductal and lobular carcinoma; expensive;	40–60 min
CT	To determine and image distant metastasis in a single exam	91%	93%	Low sensitivity; radiation risks; expensive scanner;	5 min
PET	Functional imaging of biological processes. To image metastasis or response to therapy	61.0%	80.0%	Ionizing radiation, radioactive tracer injection	90–240 min

* Sensitivity and specificity are related to the types of cancer and breast composition.

2.4 Diagnosis

After making the detection comes the diagnoses, cancer staging is the step that determines the type and extent of treatment of a patient. An essential part of this staging process is an assessment of the nodal status. Axillary lymph node (ALN) status is still one of the most important prognostic factors for patients with breast cancer. As a result, an accurate and standardised determination of metastatic nodal burden within the ALN is of utmost importance when selecting the correct treatment, thus avoiding over- as well as under-treatment.[6]

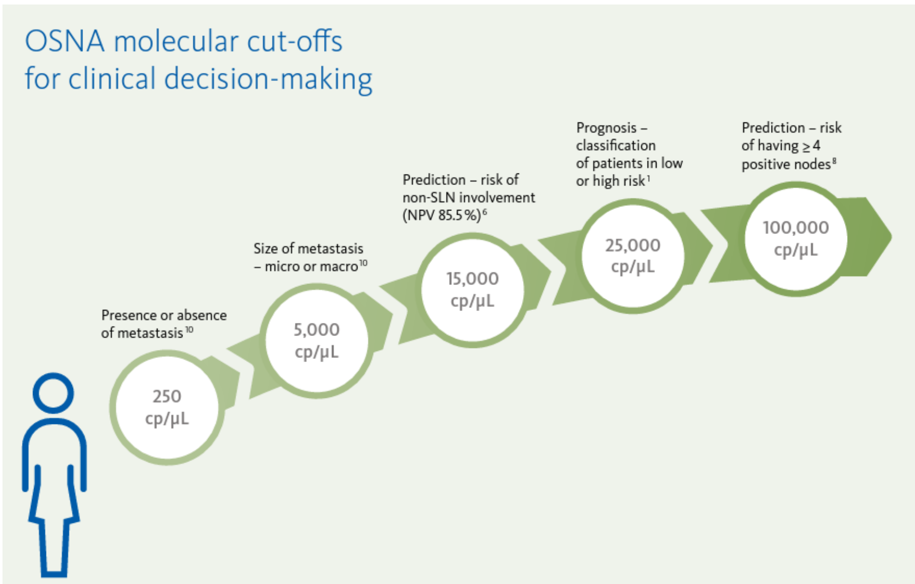

Fig. 1. Conventional breast screening methods and their limitations.[6]

2.5 Treatment

Every treatment order is individual and dependent on the cancer type and stage. Gene alterations play a crucial role during the onset and development of breast cancer. Therefore, reliable assessment of the tumour's mutational status is decisive not only in selecting cancer patients eligible for targeted therapy, but also later in monitoring response to treatment and disease progression[6]. Nowadays, a range of medications targeting tumour-specific genetic alterations are available and can be prescribed instead of, or in combination with standard chemo-therapies, immunotherapies and radio-therapies.[6]

2.6 Existing techniques and tools

In this Chapter we are going to approach the existing tools and techniques. The DSS's are appearing more and more in the heath-care subject to help the professionals make decisions with a tool that guides them in a correct path but we also have tools to help patient to have a insight of his disease.

Oncotype IQ is a tool that helps the patient understand what type o state of the disease is and help understand what are the next steps. Gives information of what stage the cancer is, if needs chemo, surgery, if is an aggressive disease. Its a tool that answers questions to patient relating to his disease and this tool uses the method OncoType DX.

The Oncotype DX assay is a 21-gene assay that predicts the likelihood of chemotherapy benefit and 10-year risk of distant recurrence to inform adjuvant treatment decisions in certain women with early-stage invasive breast cancer.[7] Informs treatment decisions and improves confidence in treatment decisions for physicians and patients. Studies suggest that the results of the assay change breast cancer treatment decisions in approximately 30% of cases.[7] Offers expanded clinical utility-Enhanced reports now include quantitative ER, PR and HER2 values.[7]

The DESIREE project aims to provide a web-based software ecosystem for the personalized, collaborative and multidisciplinary management of Primary Breast Cancer by specialized Breast Units, from diagnosis to therapy and follow-up.

DESIREE provide decision support on the available therapy options by incorporating experience from previous cases and outcomes, and thus, going beyond the limitations of existing guideline-based decision support systems (DSS). The DSS will be based on a knowledge model that will evolve with experience. Patients' cases will be represented using a novel complex Digital Breast Cancer Patient (DBCP) model, which incorporates information about the patient clinical history and diagnostic and therapeutic procedures in cycles that may last for years.[8]

The creation of a DBCP-based advanced knowledge model that incorporates clinical guidelines, clinical experience and important patient context information will provide timely advice on decisions and will reduce the number of decisions that the system is not able to reflect. It will also provide the ability to learn from experience and to evaluate the success or failure of previous decisions. It will exploit the information available both from the current case and from previous similar cases obtained by comparison using the DBCP model.[8]

Residual Cancer Burden Calculator is tool that according the pathological parameters of the patient can calculate residual cancer burden.

The following parameters are required from pathological examination in order to calculate Residual Cancer Burden (RCB) after neoadjuvant treatment[9]:

- The largest two dimensions (mms) of the residual tumor bed in the breast (largest tumor bed if multicentric disease).
- Submission of the entire largest cross-sectional area of the residual tumor bed for histologic mapping, with specific identification of those slides in the pathology report (e.g. "the largest cross-sectional area of primary tumor bed was submitted in cassettes A5 - A9")

- Histologic assessment of the percentage of the tumor bed area that contains carcinoma (all carcinoma, i.e. invasive and in situ)
- Histologic estimate of the percentage of the carcinoma in the tumor bed that is in situ
- The number of positive (metastatic) lymph nodes
- The largest diameter (mm) of the largest nodal metastasis

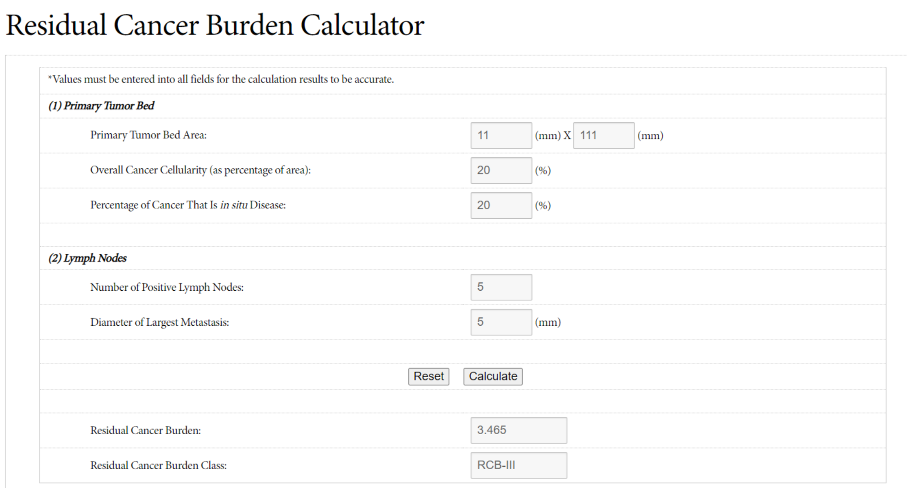

Fig. 2. Residual Cancer Burden Calculator[9]

There are also other types of technology's of DSS for breast cancer but are more focus on prediction of lifespan of the patient for different treatments options.

3 Proposal

Clinical diagnoses are the primary source of information for choosing medical treatments. Although they are of central importance in evidence-based medicine, there is a lack of usable information systems that provide data analysis and decision support capabilities for aggregated clinical diagnostic outcomes.This is partially caused by the unavailability of diagnostic data in a structured format suitable for reanalysis, and of a complete data model for aggregate-level outcomes.

We have developed a unifying data model that allows the development of evidence-based decision support in the absence of a complete data model. Our decision support system will be developed to assist healthcare professionals in identifying the tumor type. This identification depends on several factors such as the specialty of the physician the years of experience and is susceptible to human error. The goal is that the application will assist doctors in their decisions since it contains a lot of information about diagnoses and can also be improved by inserting new diagnoses.

Our data mining process will be based on a data set [10] obtained online that contains 570 samples and 30 variables that distinguish the different tumors.

Table 2. Example of factors used from data set choose.

Factors			
Radius Mean	Texture Mean	Perimeter Mean	Area Mean
Compactness Mean	Concavity Mean	Concave Points Mean	Symmetry Mean
Radius Se	Texture Se	Perimeter Se	Area Se
Compactness Se	Concavity Se	Concave Points Se	Symmetry Se
Radius Worst	Texture Worst	Perimeter Worst	Area Worst
Compactness Worst	Concavity Worst	Concave Points Worst	Symmetry Worst
Smoothness Mean	Fractal Dimension Mean	Smoothness Se	Fractal Dimension Se

Initially we will explore the data and identify the most important factors in distinguishing between a benign and malignant tumor. This exploration will be done using Data Mining techniques. After identifying the factors that have the most influence, we will also remove the average value in which they start to be problematic. This analysis is about a mammogram that gives us the values of the factors mentioned previously.

Subsequently we will create a decision tree that will serve as an aid to health professionals. This decision tree will ask simple questions to the user and with the answers obtained will indicate if it is a benign tumor, a malignant tumor or if it could not reach any conclusion. It is up to the health professionals to make the final decision.

One possible feature could be to train the model with the test data that is being entered. This would allow professionals to later input information on whether the tumor was actually benign or malignant.

Improving the model with current data and increasing the number of data in the data set would allow for a more robust model with better predictive ability

3.1 Functional Requirements

- Data mining the data set;
- Build decision tree to assist the user;
- Input information about a patient;
- Return answer to the health professional;
- Allows you to enter the tumor type of a patient tested by the decision tree;

3.2 Non-functional requirements

- Security, all data input must be protected because it is personal data;
- Performance, the system must be able to be used by multiple users;
- Availability, the system must operate for 24 hours as it will be used by multiple physicians throughout the day;
- Integration, the system must be able to communicate with current hospital computer systems;

3.3 Conceptual Architecture

Our project is a web application, which will be integrated into existing health systems in hospitals.

In the back-end part of the application, machine learning techniques are implemented on the data set to create a diagnostic model.

After this model is created, we will generate a decision tree with rules based on the model created previously, using drools [11].

Finally, this decision tree will be presented in a user interface to the end user, in which the user is asked about the most important factors for identifying the type of tumor.

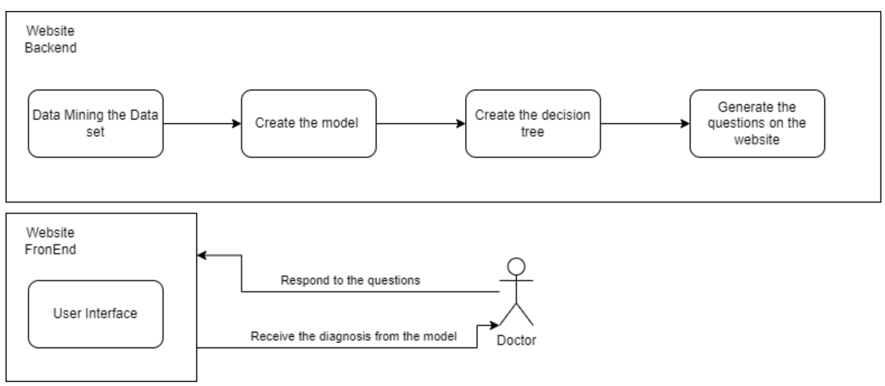

Fig. 3. Conceptual Architecture.

4 Conclusion

We can conclude that the developed project is an asset because it allows creating a history of previous diagnoses, and the improvement of the diagnosis model. After a few years, it is expected that this model will have better forecasting capabilities than health professionals specialized in the area, as it will be fed by several professionals and several hospitals. One of the next steps in this project is the sharing of information from the various hospitals. Initially at a national level and then perhaps at a global level, deep learning models are models that the larger the data set, the greater the accuracy of the model.

A moral question that we can raise is to what extent these models cannot fully replace health professionals in the near future, as they will have access to a lot of information and if the data sets are maintained/updated, it is expected that they will be able to reach around 99 % of diagnoses.

In conclusion, decision support systems are very useful in many areas, especially in medicine, where human error or misinterpretation of results can put human life at risk. These systems are especially important in critical systems or complex systems that may affect multiple people worldwide.

References

1. JiangX, WellsA, BrufskyA, Neapolitan R(2019) A clinical decision support system learned from data to personalize treatment recommendations towards preventing breast cancer metastasis. PLoSONE 14 https://doi.org/10.1371/journal.pone.0213292
2. Magnuska, Z.A.; Theek, B.;Darguzyte, M.; Palmowski, M.;Stickeler, E.; Schulz, V.; Kießling, F.Influence of the Computer-Aided Decision Support System Design on Ultrasound-Based Breast Cancer Classification.Cancers 2022,14, 277 https://doi.org/710.3390/cancers14020277
3. Massafra R,Latorre A,Fanizzi A,Bellotti R,Didonna V,Giotta F,La Forgia D,Nardone A,Pastena M,Ressa CM, Rinaldi L,Russo AOM,Tamborra P,Tangaro S,Zito A and Lorusso V (2021) A Clinical Decision Support System for Predicting Invasive Breast Cancer Recurrence:Preliminary Results.Front. Oncol.11:576007. https://doi.org/10.3389/fonc.2021.576007
4. breastcancer.org October 24, 2022, https://www.breastcancer.org/facts-statistics?gclid=CjwKCAjwtp2bBhAGEiwAOZZTuGWQDYRLk1tf6f-bsBr_SDF1GCqZ0PdQb2r0750YFsRvgBlokA5FSRoCGxsQAvD_BwE
5. Lulu Wang, Early Diagnosis of Breast Cancer, Published: 5 July 2017 https://doi.org/10.3390/s17071572
6. Sysmex, November 6, 2022, https://www.sysmex-europe.com/
7. Ellis, P.G.; Brufsky, A.M.; Beriwal, S.; Lokay, K.G.; Benson, H.O.; McCutcheon, S.B.; Krebs, M. PathwaysClinical Decision Support for Appropriate Use of Key Biomarkers.J. Oncol. Pract.2016,12, e681–e687.[CrossRef]
8. Larburu, N.; Muro, N.; Macía, I. Augmenting Guideline-based CDSS with Experts' Knowledge.In Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies, Volume 5: HEALTHINF, Porto, Portugal, 21–23 February 2017; pp. 370–376. [CrossRef]
9. Yau, C.; van der Noordaa, M.; Wei, J.; Osdoit, M.; Reyal, F.; Hamy, A.S.; Lae, M.; Martin, M.; del Monte, M.;I-SPY TRIAL Consortium; et al. Residual cancer burden after neoadjuvant therapy and long-term survivaloutcomes in breast cancer: A multi-center pooled analysis. In Proceedings of the 2019 San Antonio BreastCancer Symposium, San Antonio, TX, USA, 12–13 December 2019
10. Breast Cancer Wisconsin (Diagnostic) Data Set (2016). https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data?select=data.csv

11. Drools website (2022) https://www.drools.org/