Detection of breast cancer using skin surface electropotentials-evaluation using clinical study and data mining techniques

Vinitha Sree Subbhuraam

2010

Vinitha, S. S. (2010). Detection of breast cancer using skin surface
electropotentials-evaluation using clinical study and data mining techniques. Doctoral
thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/47158>

<https://doi.org/10.32657/10356/47158>

Nanyang Technological University

Downloaded on 26 Apr 2023 16:58:26 SGT

78
2X
2/Blue

**DETECTION OF BREAST CANCER USING SKIN
SURFACE ELECTROPOTENTIALS - EVALUATION
USING CLINICAL STUDY AND DATA MINING
TECHNIQUES**

VINITHA SREE SUBBHURAAM

School of Mechanical and Aerospace Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2010

ACKNOWLEDGEMENTS

I express my sincere gratitude to my supervisor, Assoc. Prof. Ng Yin Kwee, for his invaluable guidance in conducting this research. His constant support and sustained encouragement have been a source of inspiration for me throughout the work.

I am also grateful for the time, advice, and guidance of Dr. Rajendra Acharya, a visiting faculty in Ngee Ann Polytechnic. I am indebted to my supervisor and Dr. Rajendra for providing me with excellent guidance and support for publishing journal papers and book chapters. It has been a truly enriching learning experience.

My special thanks are due to Dr. Gregory Kaw and Dr. Chong Bee Kiang of TTSH, and Dr. Rajendra and my supervisor, for their help and involvement in establishing the clinical study protocol and getting the ethics committee approval. Thanks are also due to the doctors and radiographers of the Breast Assessment Centre of TTSH for their help and immense enthusiasm in recruiting and assisting patients for the clinical study.

I would like to thank the patients who provided me with the data needed for my work.

I am very grateful to NTU for the research scholarship, without which this work would not have been possible. The research facilities at the Advanced Design and Modeling Laboratory were used for this work, and I acknowledge and thank the staff there for providing me with a conducive research environment.

I also owe my warm thanks to my parents, Dr. Subbhuraam and Dr. Chandrakanthi, and my sister, Sinduja, for motivating me to embark on this interesting work, and for providing their love and support throughout the research period. Thanks are also due to my friends and peer researchers for their valuable support. Finally, I thank the Almighty for giving me this opportunity which I would cherish forever.

TABLE OF CONTENTS

ACKNOWLEDGEMENTSii

LIST OF FIGURESvii

LIST OF TABLESviii

LIST OF ABBREVIATIONSx

PUBLICATIONS FROM THIS THESISxiii

ABSTRACT.....xv

CHAPTER 1: INTRODUCTION 1

1.1 Background and motivation..... 1

1.2 Objectives 3

1.3 Research methodology..... 3

CHAPTER 2: LITERATURE REVIEW 6

2.1 Anatomy of the female breast 6

2.2 Breast cancer 8

2.2.1 Origin and spread of breast cancer 8

2.2.2 Risk factors 9

2.2.3 Numerical staging of breast cancer..... 10

2.3 Clinical breast imaging modalities 11

2.3.1 Breast Self Examination (BSE) 11

2.3.2 Clinical Breast Examination (CBE)..... 12

2.3.3 Mammography (MG)..... 12

2.3.4 Breast ultrasound 15

2.3.5 Computed Tomography (CT) 17

2.3.6 Breast MRI..... 19

2.3.7 Positron Emission Tomography (PET)..... 22

2.3.8 Scintimammography 24

2.3.9 Biopsy 25

2.4 Experimental breast imaging modalities..... 26

2.4.1 Dedicated breast CT..... 26

2.4.2 MRI based elastography 26

2.4.3 PET-ER imaging..... 27

2.4.4 Optical imaging..... 27

2.4.5 Thermography..... 29

Table of Contents

2.4.6	Electrical characteristics based imaging techniques	30
2.4.6.1	Electrical impedance characteristics of breast tissue	30
2.4.6.2	Electropotential characteristics of breast tissue	35
2.4.6.3	Electrical impedance scanning	37
2.4.6.4	Electrical impedance tomography	47
2.4.6.5	Biofield Diagnostic System (BDS)	51
2.4.7	Other experimental modalities	61
2.5	Comparison of the various breast imaging modalities	61
2.6	Statistical analysis techniques and definitions	67
2.6.1	Performance measures: contingency table	67
2.6.2	Performance measures: AUC	67
2.6.3	Correlation coefficient	69
2.6.4	t-test	70
2.7	Data mining techniques and definitions	70
2.7.1	Data pre-processing: Handling of missing values	70
2.7.2	Data pre-processing: Outlier detection	70
2.7.3	Data pre-processing: Normality tests	71
2.7.4	Data pre-processing: Normalization	72
2.7.5	Feature selection	72
2.7.6	Classifier development and evaluation: Classifiers used	73
2.7.7	Classifier development and evaluation: Data resampling technique	75
2.7.8	Software used	76
2.8	Association rule mining based classification - a review	76
2.8.1	Algorithms for rule generation	78
2.8.2	Rule pruning techniques	80
2.8.3	Ways to classify test samples	84
2.8.4	Various associative classifiers	85
CHAPTER 3: BDS CLINICAL STUDY METHODOLOGY		87
3.1	Inclusion criteria	87
3.2	Exclusion criteria	87
3.3	Clinical study methodology	88
CHAPTER 4: STATISTICAL ANALYSIS OF THE BDS DATASETS		91
4.1	Analysis of the TTSH dataset	91
4.1.1	TTSH dataset description and pre-processing	91

Table of Contents

4.1.2	TTSH dataset: Study of effect of risk factors on cancer incidence.....	92
4.1.3	TTSH dataset: Study of effect of lesion parameters on cancer incidence	93
4.1.4	TTSH dataset: Results of diagnostic tests.....	95
4.1.5	TTSH dataset: Biopsy reduction.....	99
4.2	Analysis of the US dataset	99
4.2.1	US dataset description and pre-processing	99
4.2.2	US dataset: Study of effect of risk factors on cancer incidence	102
4.2.3	US dataset: Study of effect of lesion parameters on cancer incidence ...	103
4.2.4	US dataset: Results of diagnostic tests	104
4.2.5	US dataset: Biopsy reduction.....	107
4.3	Discussion on results obtained using both datasets	108
CHAPTER 5: EVALUATION OF THE DATA MINING FRAMEWORK ..		111
5.1	Proposed data mining framework	111
5.2	Pre-processing.....	112
5.2.1	TTSH dataset	112
5.2.2	US dataset	113
5.3	Feature selection	114
5.3.1	TTSH dataset	115
5.3.2	US dataset	116
5.4	Classification results	117
5.4.1	TTSH dataset	117
5.4.2	US dataset	118
5.5	Discussion on results obtained using both datasets	119
CHAPTER 6: WRUBAC- DEVELOPMENT AND EVALUATION.....		123
6.1	WRUBAC algorithm	123
6.2	WRUBAC evaluation results.....	133
6.2.1	Description of the benchmark datasets	133
6.2.2	Results and discussion	136
6.3	Key features of the WRUBAC algorithm	140
CHAPTER 7: CONCLUSIONS AND FUTURE DIRECTIONS.....		142
7.1	Conclusions.....	142
7.2	Future directions	144
REFERENCES.....		146
APPENDIX A: BREAST CANCER TYPES.....		160

Table of Contents

APPENDIX B: BDS CLINICAL TRIAL FORMS 162

 B.1 Participant Information Sheet 162

 B.2 Participant Consent Form 165

 B.3 Participant Details Form 166

 B.4 BDS Sample Test Result Form 168

LIST OF FIGURES

Figure 2.1 Anatomy of the female breast.....	7
Figure 2.2 Digital mammogram: Tumor indicated by yellow circle	13
Figure 2.3 Typical ultrasound images.....	16
Figure 2.4 Sample contrast enhanced CT image - tumor indicated by arrowheads and enhancing lymph nodes indicated by arrows	19
Figure 2.5 MRI scan of the breast reveals a bright area which proved to be cancer	20
Figure 2.6 A PET image showing a focal FDG-avid left breast ovoid mass slightly superior and medial to the nipple.....	23
Figure 2.7 Optical mammogram of the left breast in mediolateral projection; the patient has multi-centric invasive ductal carcinoma as indicated by the arrows.....	29
Figure 2.8 Ideal representation of permittivity as a function of frequency	31
Figure 2.9 Breast epithelial electrical gradients.....	36
Figure 2.10 T-Scan 2000	39
Figure 2.11 Screen display in the targeted mode of T-Scan 2000.....	40
Figure 2.12 Screen display in the scanning mode of T-Scan 2000.....	40
Figure 2.13 A typical breast examination using T-Scan 2000.....	41
Figure 2.14 Mammograph Electrical Impedance Komputer	48
Figure 2.15 Data presentation screen in MEIK	49
Figure 2.16 The BDS system	52
Figure 2.17 Sample illustration of the sensor arrangement used on the patient for the BDS test, assuming that the suspicious lesion is in the upper outer quadrant of the right breast	53
Figure 2.18 Sample ROC curve.....	69
Figure 4.1 TTSH dataset: Comparison of ROC curves of ultrasound, mammography, and BDS tests	99
Figure 4.2 US dataset: Comparison of ROC curves of CBE, mammography, and BDS tests	107
Figure 5.1 Proposed data mining framework.....	112
Figure 6.1 Flow chart depicting the WRUBAC algorithm.....	133

LIST OF TABLES

Table 2.1 Summary of clinical trials carried out using T-Scan 2000	42
Table 2.2 Summary of clinical trials carried out using MEIK.....	50
Table 2.3 BDS interpretation map 1: using CMS thresholds.....	56
Table 2.4 BDS interpretation map 2: using modified thresholds	56
Table 2.5 Differences between CE and Lancet BDS devices.....	57
Table 2.6 Summary of the BDS clinical trials	60
Table 2.7 Comparison of the various breast imaging modalities	62
Table 2.8 Definition of performance measures.....	67
Table 2.9 Summary of a few associative classifiers	86
Table 4.1 TTSH dataset: Percentage breakdown of the total number of cases.....	92
Table 4.2 TTSH dataset: Age distribution and respective diagnoses	93
Table 4.3 TTSH dataset: Location of lesions	94
Table 4.4 TTSH dataset: Patient distribution in terms of diagnostic tests undertaken ..	95
Table 4.5 TTSH dataset: Ultrasound results (N=182)	96
Table 4.6 TTSH dataset: Mammography results (N=182)	97
Table 4.7 TTSH dataset: BDS results (N=182)	97
Table 4.8 TTSH dataset: Summary of ROC results of ultrasound, mammography, and BDS tests.....	98
Table 4.9 Comparison of electropotential differentials before outlier detection	102
Table 4.10 Comparison of electropotential differentials after outlier detection.....	102
Table 4.11 US dataset: Age distribution and respective diagnoses	103
Table 4.12 US dataset: Location of lesions	104
Table 4.13 US dataset: Patient distribution in terms of diagnostic tests undertaken ..	104
Table 4.14 US dataset: CBE results (N=241)	105
Table 4.15 US dataset: Mammography results (N=241)	106
Table 4.16 US dataset: BDS results (N=241)	106
Table 4.17 US dataset: Summary of ROC results of CBE, mammography, and BDS tests	107
Table 4.18 Comparison of TTSH and US datasets.....	110
Table 5.1 Derived electropotential based features.....	114
Table 5.2 TTSH dataset: Features selected using filter and wrapper methods.....	116

List of Tables

Table 5.3 US dataset: Top five features according to Chi-Squared method and the ranks of these features according to GR, IG, and SymU filters.....	116
Table 5.4 US dataset: Features selected using filter and wrapper methods.....	117
Table 5.5 TTSH dataset: Performance measures of the five classifiers	118
Table 5.6 US dataset: Performance measures of the five classifiers	119
Table 6.1 Description of the benchmark datasets	134
Table 6.2 WRUBAC algorithm evaluation results	137
Table 6.3 IRIS dataset: non-updated and updated rules	139
Table 6.4 Comparison of WRUBAC, BPNN, SVM, and LDA/WFS classifier results for the US dataset.....	140

LIST OF ABBREVIATIONS

AC	Asymptomatic Central sensor
AD	Asymptomatic Diagonal sensor
AH	Asymptomatic Horizontal sensor
AI	Asymptomatic Inner sensor
AL	Asymptomatic Lower sensor
AO	Asymptomatic Outer sensor
ARM	Association Rule Mining
AU	Asymptomatic Upper sensor
AUC	Area under the ROC Curve
AV	Asymptomatic Vertical sensor
BCDD	Breast Cancer Detection Device
BDS	Biofield Diagnostic System
BI-RADS	Breast Imaging Reporting and Data System
BPNN	Back Propagation Neural Network
BSE	Breast Self Examination
CAD	Computer Aided Diagnostic
CBA	Classification Based on Associations
CBE	Clinical Breast Examination
ce-CT	contrast-enhanced CT
CFS	Correlation-based Feature Selection
CMS	Clinical Management System
CS	Chi-Squared
CT	Computed Tomography
DCE-MRI	Dynamic Contrast Enhanced-MRI
DCIS	Ductal Carcinoma in Situ
DMIST	Digital Mammographic Imaging Screening Trial
DOR	Diffuse Optical Imaging
DOS	Diffuse Optical Spectroscopy
DOT	Diffuse Optical Tomography
ECG	Electrocardiogram
EIM	Electrical Impedance Mammogram
EIS	Electrical Impedance Scanning

List of Abbreviations

EIT	Electrical Impedance Tomography
EPD	Electropotential Differentials
ER	Estrogen Receptor
ERT	Estrogen Replacement Therapy
FDA	Food and Drug Administration
FDG	Fluorodeoxyglucose
FES	[¹⁸ F]-16 α -[fluororestradiol]
FFDM	Full Field Digital Mammography
FFS	Filter Feature Subset
FIS	Fuzzy Inference System
FPR	False Positive Rate
FP-tree	Frequent Pattern-tree
FSM	Film-Screen Mammography
GR	Gain Ratio
HRT	Hormone Replacement Therapy
IDC	Invasive Ductal Carcinoma
ILC	Invasive Lobular Carcinoma
IR	Information Gain
LCIS	Lobular Carcinoma in Situ
LDA	Linear Discriminant Analysis
LDA/WFS	LDA classifier developed using WFS feature subset
LHS	Left-Hand Side
LOS	Level of Suspicion
MCC	Microcalcification Clusters
MEIK	Mammograph Electrical Impedance Komputer
MG	Mammography
MRI	Magnetic Resonance Imaging
MRM	Magnetic Resonance Mammography
MRSI	Magnetic Resonance Spectroscopy Imaging
NC	Normalization Constant
NIR	Near Infrared
OCP	Oral Contraceptive Pills
PEM	Positron Emission Mammography

List of Abbreviations

PER	Pessimistic Error Rate
PET	Positron Emission Tomography
PMS	Post Market Surveillance
PNN	Probabilistic Neural Network
RBF	Radial Basis Function
RF	Radio Frequency
RGT	Rule Goodness Threshold
RHS	Right-Hand Side
ROC	Receiver Operating Characteristic
ROI	Region of Interest
RUS	Rule Updation System
SC	Symptomatic Central sensor
SD	Symptomatic Diagonal sensor
SH	Symptomatic Horizontal sensor
SI	Symptomatic Inner sensor
SL	Symptomatic Lower sensor
SO	Symptomatic Outer sensor
SPECT	Single Photon Emission Computed Tomography
SU	Symptomatic Upper sensor
SV	Symptomatic Vertical sensor
SVM	Support Vector Machine
SymU	Symmetrical Uncertainty
TEP	Transepithelial Electrical Potential
TTSH	Tan Tock Seng Hospital
US	United States
WAR-BC	Weighted Association Rule Based Classification
WFS	Wrapper Feature Subset
WOBC	Breast Cancer Wisconsin (Original)
WRUBAC	Weighted Rule Updation Based Associative Classification

PUBLICATIONS FROM THIS THESIS

Journal papers

Ng, E.Y.-K., Vinitha Sree, S., Ng, K.H., Kaw, G. (2008). The use of tissue electrical characteristics for breast cancer detection: A perspective review. *Technology in Cancer Research and Treatment*, 7(4), 295-308.

Vinitha Sree, S., Ng, E.Y.-K., Kaw, G., Rajendra Acharya, U., Chong, B.K. (2009). The use of skin surface electropotentials for breast cancer detection: Preliminary clinical trial results obtained using the Biofield Diagnostic System. *Journal of Medical Systems*, DOI 10.1007/s10916-009-9343-0.

Vinitha Sree, S., Ng, E.Y.-K., Rajendra Acharya, U. (2010). Data mining approach to evaluating the use of skin surface electropotentials for breast cancer detection. *Technology in Cancer Research and Treatment*, 9(1), 95-106.

Vinitha Sree, S., Ng, E.Y.-K., Kaw, G., Rajendra Acharya, U., Chong, B.K. (2010). Evaluation of the efficiency of Biofield Diagnostic System in breast cancer detection using clinical study results and classifiers. *Journal of Medical Systems*, DOI 10.1007/s10916-010-9441-z.

Vinitha Sree, S., Ng, E.Y.-K., Rajendra Acharya, U., Tan, W. (2010). Breast imaging systems: a review. *Journal of Mechanics in Medicine and Biology*, 10(1), 5-34.

Conference papers

Vinitha Sree, S., Ng, E.Y.-K., Kaw, G., Rajendra Acharya, U. (2009). Use of data mining techniques for improved detection of breast cancer with Biofield Diagnostic System. *11th International Conference on Engineering Applications of Neural Networks (EANN 2009)*, proceedings published in Springer's *Communications in Computer and Information Science series*, 43, 444 - 452.

Book chapters

Vinitha Sree, S., Ng, E.Y.-K., Ng, K.H., Rajendra Acharya, U., Lim Choo Min, Suri, J.S. (2009). Perspective review of the use of electrical impedance and biopotential based devices as possible adjunct breast diagnostic tools. In E.Y.-K. Ng, R.U. Acharya, J.S. Suri (Eds.), *Distributed Diagnosis and Home Healthcare (D2H2)*. USA/Canada: American Scientific Publishers, ISBN: 1-58883-158-2.

Tan, W., Vinitha Sree, S., Rajendra Acharya, U., Ng, E.Y.-K., Yu, W., Suri, J.S. (2010). The human breast. In E.Y.-K. Ng, R.U. Acharya, J.S. Suri (Eds.), *Performance evaluation of breast cancer - screening, diagnosis and treatment*. USA/Canada: American Scientific Publishers, ISBN: 1-58883-156-6 (in press).

Vinitha Sree, S., Ng, E.Y.-K., Chong, B.K., Rajendra Acharya, U., Kaw, G. (2010). Clinical assessment of breast lesions using skin surface electropotentials: Preliminary

Publications from this thesis

results of a prospective clinical trial of the Biofield Diagnostic System. In E.Y.-K. Ng, R.U. Acharya, J.S. Suri (Eds.), *Performance evaluation of breast cancer - screening, diagnosis and treatment*. USA/Canada: American Scientific Publishers, ISBN: 1-58883-156-6 (in press).

ABSTRACT

Breast cancer is the most frequently occurring malignancy in women, and the mortality rate is high compared to other cancers. For better survival odds and lesser use of treatments and therapies and, therefore, lesser side-effects, many imaging systems are continually being developed to diagnose this disease as early as possible. In this thesis, a comprehensive literature review of the various breast imaging systems, which are currently being used clinically or studied experimentally, is given. A perspective review of the electrical characteristics of normal and cancerous breast tissue, followed by a description of the various modalities that use these characteristics for breast cancer detection is presented.

The Biofield Diagnostic System (BDS) is an adjunct breast cancer diagnostic modality that uses skin surface electropotentials to differentiate benign from malignant breast lesions. It works on the principle that the irregular rapid cell division associated with cancer causes the area around the cancer to be more depolarized than the nearby normal tissue, thereby resulting in large electropotential differences between the cancerous and normal areas of the breast. These differentials are transmitted to the breast surface by means of volume conduction and are measured by BDS using single-use sensors. One of the main objectives of this work was to evaluate the efficiency of BDS in a clinical setting. For this purpose, a clinical study was designed and carried out in collaboration with physicians from Tan Tock Seng Hospital (TTSH), Singapore. 182 women volunteered for the BDS examination. The test demonstrated high values for sensitivity (96.23%), specificity (93.80%), and accuracy (94.51%). Also, BDS presented a 93.18% accuracy level in predicting unwanted biopsies.

Another objective was to study the utility of data mining techniques in improving the disease detection accuracy of BDS. For this part of the work, two BDS datasets were used - the dataset collected in the TTSH clinical study, and another US dataset that was obtained directly from the BDS manufacturer. A generalized classification framework is proposed. It comprises of (a) a data pre-processing element, (b) a feature selection element, (c) a classifier development element, and (d) a classifier evaluation element. This framework can be employed to determine the *feature subset - classifier combination* that gives the best prediction accuracy for any analyzed dataset. A

Abstract

multitude of supervised learning based classifiers such as Back Propagation Neural Network (BPNN), Probabilistic Neural Network (PNN), Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), and a fuzzy inference classifier were tested. It was observed that the LDA classifier, using the feature subset selected via wrapper feature selection technique, significantly improved the sensitivity and accuracy of the US dataset. Also, this feature subset reduced the mild subjective interpretation associated with the current BDS prediction methodology, thereby opening up new development avenues for the BDS device.

A novel associative classification algorithm, Weighted Rule Updation Based Associative Classification (WRUBAC), was developed in this work. WRUBAC is a type of associative classification algorithm that incorporates the traditional three steps, such as rule generation, rule pruning, and classification, followed in developing an associative classifier. Apart from these steps, the algorithm also has additional rule pruning steps and a novel rule updation process wherein the weight of each rule is updated based on the training dataset. WRUBAC was tested on six benchmark datasets and the two BDS datasets. Better or similar classification accuracies were obtained with the added advantages of lesser space and computational time requirements, and lesser number of rules when compared to the already available associative classifiers.

The black-box approach of neural networks reduces the possibility of easy interpretation of the results. However, because of the use of rules for classification, the WRUBAC algorithm has a better interpretability. Hence, the algorithm has good scope in the field of medical diagnostics where doctors prefer the information, which is used by a diagnostic modality to predict the presence or absence of a disease, to be transparent and comprehensible. Using WRUBAC, the domain knowledge can be extracted in the form of rules, and the rules can be used to develop a graphical user interface based classification tool that could improve the prediction accuracy of diagnostic devices.

CHAPTER 1: INTRODUCTION

1.1 Background and motivation

The United States National Cancer Institute has estimated that one in eight women will be diagnosed with breast cancer at some point of time in their lives (Horner et al., 2009). Breast cancer is the second leading cause of cancer related death in women. It has been estimated that 1,92,370 women would be diagnosed with invasive cancer, 62,280 women with ductal carcinoma in situ, and 40,170 women would die of breast cancer in 2009 in the United States (Horner et al., 2009). It is also the most common cancer affecting women in Singapore, with around 1000 women diagnosed every year (Singapore cancer statistics, 2009). In spite of advances in the diagnosis and treatment of breast cancer, it still continues to be a predominant cause of death among women all around the world. Because of the high mortality rate and lack of proven prevention strategies and cure for this disease, early detection is the only ray of hope for having a control over the pain, emotional impact, and possible death associated with this disease. Hence, as much as the scientists are trying for a possible cure, an equal effort is also going into developing efficient standalone or adjunct early detection modalities.

There are many breast imaging systems available in the market to facilitate diagnosis of breast cancer. Some of them are used for screening, some are used for diagnosis, and most others are used for adjunctive evaluation. Mammography (MG) is currently the most widely used technique for breast cancer detection. Even though it has demonstrated good performance and fast patient throughput in older women, it is not effective in detecting cancers in younger women (≤ 40 years of age) who have greater breast density. In a study conducted by Carney et al. (2003), MG recorded a sensitivity of 87% and specificity of 96.9% for women with fatty breasts. However, for younger women with extremely dense breasts, the sensitivity of MG was only 62.9%, and the specificity was 89.1%. This implies that the sensitivity increases with age and with less denser breasts. Moreover, if MG is used at such an early age, there is a risk of increased radiation exposure. As a result, for younger women, generally, breast self examination, Clinical Breast Examination (CBE), and ultrasound are the recommended procedures. By the time CBE detects a cancer, the cancer is usually at an advanced stage. Breast ultrasound is frequently used to evaluate breast

abnormalities that are found with screening or diagnostic MG or CBE. In spite of the many advantages of this procedure, it is not as efficient in imaging microcalcifications as MG. Microcalcifications are often the first indication of breast cancer. Breast Magnetic Resonance Imaging (MRI) is useful when women have implants or suspected multifocal carcinoma (Schnall, 2003). Though it is highly sensitive, especially for high-risk women, the specificity is low and it is more time consuming and expensive. Thus, most of the currently available diagnostic and screening modalities either fail at times to detect the abnormality or give ambiguous results or are only efficient in a certain subgroup of patients. Therefore, in many cases, doctors recommend biopsies, and about 80% of the time, the biopsies turn out to be negative (Burns, 1997). Moreover, such unwanted biopsies cause emotional trauma for the patient, and also result in increased healthcare expenditure. Thus, there is a need for an adjunct modality to reassure the doctors of their diagnostic predictions and also to reduce the number of unwanted biopsies.

Several techniques such as MRI-based elastography, microwave imaging, thermography, optical mammography, electrical impedance tomography, electrical impedance scanning, and electropotential based technique are currently being researched for the efficient detection of potentially curable early stage breast cancers. Since the 1920s, many research studies have concluded that malignant breast tumors have significantly different electrical impedance or electropotential values compared to those of normal tissues and benign lesions. The Biofield Diagnostic System (BDS) is an adjunct breast cancer diagnostic modality that uses skin surface electropotentials to differentiate benign from malignant breast lesions. It works on the principle that the irregular rapid cell division associated with cancer causes the area around the cancer to be more depolarized than the nearby normal tissue, thereby resulting in large electropotential differences between the cancerous and normal areas of the breast. These differentials are transmitted to the breast surface by means of volume conduction and are measured by BDS using single-use sensors. The BDS test is non-invasive, radiation-free, cost-effective, and does not cause discomfort to the patient. It can be repeated as often as needed. The test results obtained are objective and acquired within a maximum of 15 minutes. Therefore, BDS was chosen for further evaluation in this work.

Since cancer is a complex disease with a varied pathology, accomplishment of early detection is difficult with the available diagnostic standalone or adjunct devices. There is a general consensus that mining of the data or images obtained from these devices would present new and highly significant predictive information that might result in more accurate disease detection. Thus, it is evident that data mining approaches can be used to support the current diagnostic devices to improve their cancer detection accuracy. Therefore, in this work, the efficiency of BDS in breast cancer detection is evaluated using both clinical study and data mining techniques.

1.2 Objectives

The following are the key objectives of this work.

- i. To evaluate the efficiency of the Biofield Diagnostic System (BDS) in a clinical setting.
- ii. To study the utility of data mining techniques in improving the disease detection accuracy of BDS.
- iii. To develop a novel associative rule mining based classifier that is relatively more transparent and comprehensible when compared to other classifiers.

1.3 Research methodology

The research methodology adopted in this work is outlined below.

Literature review

To establish the need for better adjunct breast cancer detection modalities, a review of the various breast imaging systems, which are currently being used clinically or studied experimentally, was done. In the process, a comprehensive perspective review of the electrical characteristics of normal and cancerous breast tissue, and the various modalities that use these characteristics for breast cancer detection, was carried out. In addition, a study of the anatomy of the female breast, types, stages, and risk factors pertaining to breast cancer, the statistical techniques used in the work, the basic steps employed in any data mining methodology, and a review of associative classifiers was also carried out. These reviews are presented in *Chapter 2*.

BDS clinical study - patient recruitment and examination

The patients were recruited at the Tan Tock Seng Hospital (TTSH), Singapore, based on the inclusion and exclusion criteria stipulated for the BDS examination. Information regarding the patient's demographic parameters was recorded on a patient details form. Informed consent was obtained from the patients. The details of the clinical study methodology are given in *Chapter 3*.

Statistical analysis of the BDS datasets

Two BDS datasets were available for further analysis. One was the clinical study dataset, henceforth referred to as the *TTSH dataset*. Another BDS dataset, referred as the *US dataset*, was obtained from the manufacturer. The effect of demographic parameters such as age, menstrual history, pregnancy history, and family history of cancer, and the effect of lesion based parameters such as size, location, and palpability status on the cancer incidence were studied. The sensitivity, specificity, accuracy, and Area under the Receiver Operating Characteristic Curve (AUC) measures were determined to quantify the efficiency of BDS. These results are presented in *Chapter 4*.

Use of a data mining framework on the BDS datasets

To study the use of data mining techniques, a generalized data mining framework was developed. This framework can be employed to determine the feature subset- classifier combination that gives the best prediction accuracy for any analyzed dataset. Based on this framework, a multitude of supervised learning based classifiers such as Back Propagation Neural Network (BPNN), Probabilistic Neural Network (PNN), Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), and a fuzzy inference classifier were tested using the two BDS datasets. The results and key observations are discussed in *Chapter 5*.

Development and evaluation of WRUBAC

In the course of search for more transparent classifiers, a novel classification algorithm, Weighted Rule Updation Based Associative Classification (WRUBAC), which is based on the concept of Association Rule Mining (ARM), was then developed in this work. WRUBAC is a type of associative classification algorithm that incorporates the traditional three steps, such as rule generation, rule pruning, and

classification, followed in developing an associative classifier. Apart from these steps, the algorithm also has additional rule pruning steps and a novel rule updation process wherein the weight of each rule is updated based on the training dataset. WRUBAC was then evaluated using six benchmark datasets and the two BDS datasets. The algorithm, evaluation results, and novel features of WRUBAC are presented in *Chapter 6*.

Conclusions and future directions

Conclusions are drawn, and the possible future research directions for BDS and the WRUBAC algorithm are presented in *Chapter 7*.

CHAPTER 2: LITERATURE REVIEW

This chapter presents the literature study done on the anatomy of the female breast, types, stages, and risk factors of breast cancer, and the various breast imaging systems that are currently being used clinically or studied experimentally. A comprehensive perspective review of the electrical characteristics of normal and cancerous breast tissue, and the various modalities that use these characteristics for breast cancer detection, is also presented. Moreover, details about the statistical techniques used in the work, the basic steps employed in any data mining methodology, and a review of associative classifiers are given.

2.1 Anatomy of the female breast

The breasts, anatomically known as the mammary glands, are distinctive features exclusive to mammals, such as humans. Both genders can have mammary glands. However, in the case of human species, when puberty is reached, they become more expressive in women than in men. They are glands, more of sebaceous type, which evolved to become milk-producing organs to provide nourishment for the newborn offspring. The attached base of the breast extends vertically from the second or third to the sixth rib, and in the transverse plane, from the sternal edge medially, almost to the midaxillary line laterally. The breasts are attached to the chest wall by fibrous strands called Cooper's ligaments. These ligaments provide the mobility and support to the breast.

The main parts of the breast are the nipple, lobules that produce the milk, ducts that serve as channels which transport milk from lobules to the nipple, the areola which is a pink or brown pigmented area surrounding the nipple, and fat and connective tissues which fill up the spaces that surround the lobules and ducts. The breast is composed of two main types of tissues: glandular tissues and stromal tissues. Glandular tissues house the lobules and ducts, while stromal tissues include fatty and fibrous connective tissues of the breast. The breast is also made up of lymphatic tissue that removes cellular fluids and waste. The glandular tissue is arranged in 15-20 lobes, each leading to a lactiferous duct which converges at the nipple (Figure 2.1). The lobes are divided

into multitude of lobules. Generally, breast cancers arise at the junction of these lobules. The breast structure also has connective tissue, blood vessels, and nerves.

Figure 2.1 Anatomy of the female breast
(Image courtesy of the American Cancer Society)

At maturity, the glandular portion of the breast has a protuberant conical shape. The size of the breasts varies based on many factors such as the volume of breast tissue, family history, age, degree of hormonal influence on the breast (especially estrogen and progesterone), menstrual status etc. Breast shape and appearance undergo a number of changes as a woman ages. In young women, the breast skin stretches and expands as the breasts grow, creating a rounded appearance. Young women tend to have denser breasts (more glandular tissue) than older women. During each menstrual cycle, the breast structure undergoes changes under the influence of changes in ovarian hormone levels. Around day 26 of the cycle, cell proliferation is maximal. Thereafter, the ductal system undergoes reduction, with epithelial cell apoptosis greatest on day 28. After menopause, due to the changes in hormonal development, the glandular component of the breast is replaced by fat and connective tissue. The breasts also tend to increase in size and sag because the connective tissue loses its strength.

2.2 Breast cancer

This section describes the nature of breast cancer and its types, probable risk factors, and the numerical system used for staging breast cancer.

2.2.1 Origin and spread of breast cancer

Cancer is a group of diseases characterized by abnormal, uncontrolled rapid growth of cells. Breast cancer refers to the erratic growth and proliferation of cells that originate in the breast tissue. A group of rapidly dividing cells may form a lump or mass of extra tissue. These masses are called tumors. Tumors can either be cancerous (malignant) or non-cancerous (benign). Malignant tumors penetrate and destroy healthy body tissues. A group of cells within a tumor may also break away and spread to other parts of the body. Cells that spread from one region of the body into another are called metastases.

Although breast cancer is a malignancy that originates from the parenchyma of the breast, not every breast cell has the same cell of origin. Breast cancer can originate from either the cells that line the ducts or from cells within the breast lobules. Malignant diseases develop initially on the epithelial level, indicated by increasing numbers of proliferating epithelial cells. At this point, the disease is believed to be still reversible. If it is allowed to progress further, and if it advances into carcinoma in situ, it becomes irreversible and breaks out of the duct to invade the rest of the breast tissues and spreads after gaining access to the lymphatic and blood vessel pathways.

A benign tumor behaves differently compared to a malignant tumor. Benign lumps are often caused by fibrocystic changes. Cysts are fluid-filled sacs and fibrosis refers to the formation of connective tissue or scar tissue. Fibrocystic changes can cause breast swelling and pain. The breasts may feel lumpy and sometimes there is clear or slightly cloudy nipple discharge. Though benign breast tumors are abnormal growths, they do not spread outside the breasts.

Breast cancer can be broadly classified into two categories: non-invasive and invasive. Cancer cells confined to the ducts or lobules of the breast belong to the first category. Ductal Carcinoma in Situ (DCIS) is the most common form of non-invasive breast

cancer. Lobular Carcinoma in Situ (LCIS) is less common and considered a marker for increased breast cancer risk. In invasive breast cancer, cancer cells break through the duct and lobular wall, and invade the surrounding fatty and connective tissues of the breast. Cancer can be invasive without being metastatic to the lymph nodes or other organs. There are many other forms of breast cancer, and some of them are briefly described in Appendix A.

2.2.2 Risk factors

Many risk factors for breast cancer have been cited in the literature. Some are established factors, while few others remain controversial. Listed below are the most probable risk factors for breast cancer.

- **Gender:** Simply being a woman is the main risk factor for developing breast cancer. Although breast cancer can affect men, this disease is about 100 times more common in women than in men.
- **Age:** Most women over the age of 50 have a higher risk for breast cancer.
- **Genes:** Inherited mutations in breast cancer related genes such as BRCA1 and BRCA2 may cause breast cancer in 10% of the cases. If the mutated gene has been inherited from either of the parent, the risk is high.
- **Self history:** A women with cancer in one breast has a 3- to 4- fold increased risk of developing cancer in the other breast.
- **Family history:** The risk of getting breast cancer doubles for a woman who has first-degree relatives, that is mother, sister and/or daughter, with the disease; or increases five times for a woman with two or more second-degree relatives, such as cousins or aunts, with a history of breast cancer.
- **Previous breast biopsy:** If the earlier biopsies had indicated proliferative breast disease without atypia or usual hyperplasia, the risk increases by about two-fold. The risk increases by 4 to 5 times, if the previous biopsy result indicates atypical hyperplasia.
- **Race:** White women have a slightly higher risk of developing than African-American women. In spite of this, African-Americans are more likely to die of breast cancer as they are often diagnosed at an advanced stage, that is, when breast cancer is more difficult to be treated and cured. On the other hand, Asian, Hispanic and American-Indian women have a lower risk of getting breast cancer.

-
- **Previous breast irradiation:** Women who have undergone radiation therapy in the chest area have an increased risk for breast cancer.
 - **Menstrual history:** Women who started menstruating at an early age (< 13 years) or who went through menopause at a late age (> 50 years) may have a slightly higher risk of getting breast cancer. This is because of the influence of the body's hormonal environment on the rate of developing breast cancer. Of the naturally occurring hormones, the one with the most powerful influence on breast cancer development is estrogen. Thus, the longer the body is exposed to estrogen, the higher the risk of developing breast cancer.
 - **Hormone usage:** Women who are currently using Oral Contraceptive Pills (OCP) have a slightly higher risk than women who have stopped using it more than 10 years ago. Similarly, the increased risk of breast cancer in Hormone Replacement Therapy (HRT) users reduces within 5 years of stopping HRT.
 - **None or late child-bearing:** Women who have had no children or who have had their first child after the age 30 have a slightly higher risk of breast cancer.
 - **Alcohol:** Use of alcohol is also linked to increased risk of breast cancer.
 - **Obesity:** Obesity is also associated with an increased risk of developing breast cancer, especially for women after menopause. Besides the production of estrogen by the ovaries, fat tissues also contribute to the level of estrogen by changing some other hormones into estrogens. Therefore, having more fat tissues leads to an increased level of estrogen, this in turn results in an increased breast cancer risk.
 - **High fat diet:** A high-fat diet often makes women have excess metabolic energy, coupled with low physical activity. The longer these excess metabolic energies remain in the system, the higher the possibilities of resulting in an elevated level of ovarian steroid hormones. These hormones are influential on estrogen and progesterone, which are crucial in the development of breast cancer.

Other controversial risk factors include smoking, induced abortion, breast implants, and antiperspirants.

2.2.3 Numerical staging of breast cancer

The extent of breast cancer is described by a stage 0 – IV system, which contains the T (Tumor size), N (Palpable lymph nodes), and M (Metastases) components. The numerical staging of a breast cancer describes its size and extent to which it has

spread. Whether the lymph nodes contain cancer cells is an important factor for determining treatment, and predicting survival. Hence, assessment of lymph nodes is part of the staging system. The following are the various classification stages according to this staging system:

- **Stage 0 or *in situ*:** Stage 0 cancer basically refers to a cancer that has not spread beyond the breast ductal system. DCIS and LCIS belong to this category.
- **Stage I:** The primary cancer has a maximum diameter of 2 cm and has not spread to the lymph nodes.
- **Stage IIA:** The primary tumor has a diameter in the range 2 to 5 cm and has not spread to the lymph nodes.
- **Stage IIB:** The primary tumor has a diameter between 2 and 5 cm and has spread to the axillary lymph nodes; or the primary node is over 5 cm and has not spread to the lymph nodes.
- **Stage IIIA:** Stage IIIA refers to primary breast cancers of any kind that has spread to the axillary lymph nodes and to axillary tissues.
- **Stage IIIB:** The primary breast tumor can be of any size, which has attached itself to the chest wall and has spread to the pectoral lymph nodes.
- **Stage IV:** The primary cancer has spread out of the breast to other parts of the body (such as bone, lung, liver, brain).

2.3 Clinical breast imaging modalities

This section describes in detail the various breast cancer detection modalities that are currently used in clinical practice.

2.3.1 Breast Self Examination (BSE)

BSE is a visual and manual examination of the breast that can be easily carried out by the women themselves at home. It is done by physically feeling the breast for any lump that normally does not come and go with each menstrual cycle. BSE, if done regularly, permits the detection of some early changes and small lumps in the breast. If a lump is felt, it is necessary to go for further examination.

2.3.2 Clinical Breast Examination (CBE)

A CBE is a physical examination of the breasts by a trained medical professional. It includes inspection and palpation of the entire breast area including the lymph node areas under the arm. A small percentage of breast cancers are not detected by mammography but can be felt during a CBE. Clinical breast exams are an important part of breast cancer screening and can be thought of as a complementary technique to mammography.

2.3.3 Mammography (MG)

Currently, mammography is the only U.S. Food and Drug Administration (FDA) approved screening tool to detect breast cancer in asymptomatic women (Joy et al., 2005). Mammography uses a narrow band of low keV X-rays to create detailed images (mammograms) of the breast. *Screening mammography* is an X-ray examination of the breasts in a woman who is asymptomatic. In this procedure, usually two views of the breast are taken, namely, mediolateral oblique view and cranial-caudal view. The main purpose of screening mammography is to detect small tumors at the earliest stage. If the woman has a symptomatic breast or has an abnormal screening mammogram, *diagnostic mammography* follows. Diagnostic mammography determines the exact size and location of the breast abnormality and images the surrounding tissue and lymph nodes. It includes several additional views of the breast and is hence, more time consuming compared to screening mammography.

Mammography uses low dose X-ray, high-contrast and high-resolution detectors, and an X-ray system designed specifically for imaging the breasts. In this procedure, the breast is gently compressed between two plates for a few seconds. Breast compression decreases the thickness of the breast tissue under study, and hence, improves the visibility of small lesions, reduces the radiation dose, enhances the image quality, and results in better immobilization of the breast. The compressed breast is exposed to X-rays. The attenuated X-ray photons pass through a grid, interact with an image receptor, and are then absorbed as a latent image on the recording devices. In the case of Film-Screen Mammography (FSM), the recording device is a film screen. The other clinical mammography technique is the Full Field Digital Mammography (FFDM) that

uses digital detectors as the recording media. The digital images provided by FFDM offer many advantages over its film counterpart (see Table 2.7).

Factors such as the density of the breast tissue, and the location, size and shape of the tumor play important roles in cancer detection using mammography. Cancerous masses and microcalcifications have a tendency to have a greater density than the non-cancerous masses like lipomas and cysts. The greater the density of the finding, the brighter it will appear on a mammogram. Thus, glands, connective tissue, cancerous and non-cancerous masses, microcalcifications, and fat appear as various shades of grey on a mammogram. A sample digital mammogram is shown in Figure 2.2.

Figure 2.2 Digital mammogram: Tumor indicated by yellow circle
(Schulz-Wendtland et al., 2009)

The largest study to date comparing digital and film mammography is the Digital Mammographic Imaging Screening Trial (DMIST) where 49,528 asymptomatic women were evaluated by both types of mammographic techniques (Pisano et al., 2005). The study concluded that there was no significant difference in cancer detection by either technique. However, a significantly better detection was demonstrated by the digital counterpart in subgroup analyses for women of age 50 or younger, premenopausal women, or women with dense breasts (Pisano et al., 2008). This better detection can be attributed to the improved contrast resolution of the digital

mammography. An update on digital mammography techniques can be found in Schulz-Wendtland et al. (2009).

Digital tomosynthesis mammography: In digital tomosynthesis mammography or 3D mammography, the breast is slightly compressed to keep it in a stable position during imaging. The X-ray tube moves in an arc around the breast to capture multiple low dose images at various angles. Since the breast is imaged at many possible angles, the abnormalities can be better detected. A computer is then used to process the acquired images and to construct cross-sectional 3D images of the breast. Currently, the most significant limitation of this tomographic technology is the amount of time required to reconstruct the image (Wu et al., 2003). The technique is being used in Europe and Asia, and is yet to receive FDA approval for use in the United States.

CAD in mammography: Commercial Computer Aided Diagnostic (CAD) software is available in the market to help radiologists detect masses and microcalcifications in mammograms with improved accuracy. The presence of Microcalcification Clusters (MCC) is an important sign for the detection of early breast carcinoma. Around 60-80% of the cancers show the presence of MCC on biopsies. However, these clusters are small (0.1-1 mm), have low contrast, have various shapes and sizes, and are quite similar to the surrounding tissues. Therefore, detection and classification of MCC using CAD tools can improve the accuracy of mammography to a great extent. Images are first enhanced using a variety of standard image enhancement methods mainly to sharpen the boundaries of the Region of Interest (ROI) and to increase the contrast between the ROI and the nearby normal tissue. The ROIs are then segmented through common statistical, region-based, or morphological approaches, and important features (texture-based, fractal dimension, etc.) are extracted for subsequent MCC classification. A review of literature on CAD tools and techniques for microcalcification detection can be found in Cheng et al. (2003) and Elter et al. (2009).

Another important indicator of malignancy is the presence of abnormal masses. CAD tools have been developed to detect the masses in mammograms. It has been found that it is more difficult to detect masses than MCC because masses are quite subtle, have smoother boundaries than MCC, often occur in dense areas of the breast, and have various shapes (circumscribed, speculated, lobulated, or ill-defined). Moreover,

the density of masses on mammograms blends with that of normal tissue. The series of steps, such as pre-processing, ROI segmentation, feature extraction, and classification, employed in CAD tools for mass detection are similar to those used for MCC detection. The techniques used in each of these steps are customized to mass detection. Usually shape-based, intensity-based, and texture-based features are extracted from the masses, and used in classifiers such as neural networks, decision trees, etc. A good review of literature on CAD tools and techniques for mass detection in mammograms can be found in Cheng et al. (2006) and Elter et al. (2009).

2.3.4 Breast ultrasound

Ultrasound imaging, often popularly known as sonography, is another diagnostic tool used to detect breast lesions. It has been approved by the FDA as a routine adjunctive clinical diagnostic tool. When mammographic images of breast lesions are not specific enough to identify the shape and location, sonographic imaging could be employed in combination to enhance the level of sensitivity. Since it is an adjunctive test for further examination of a palpable mass or mammographic abnormality, breast ultrasound is typically targeted at the suspicious area of the lesion. Ultrasound imaging is based on the principle that when a sound wave strikes an object, it bounces back, passes through, or gets scattered. The scattered ultrasound waves vary depending on the density and bulk compressibility of the object. In the case of breast, at diagnostic imaging frequencies (< 20 MHz), the structures that cause the strongest scattering include coarse calcium deposits followed by fibrous tissue. By measuring these scattered waves, it is possible to determine the distance, size, shape, consistency, and uniformity of the object. Thus, ultrasound is applied to detect changes in appearance and function of organs, tissues, and abnormal masses such as tumors.

In an ultrasound examination, the radiologist moves a hand-held transducer over the area of interest in the breast. The transducer directs inaudible, high-frequency sound waves into the breast tissues. The scattered sound waves are recorded by a sensitive microphone in the transducer. The recorded signals are displayed as 2D images on a computer. As the sensor is moved over the breast, continuous real-time images can be captured. Figure 2.3 illustrates typical ultrasound images of different breast conditions. The conditions are depicted by the relatively darker areas on the image.

Figure 2.3 Typical ultrasound images
(Sivaramakrishna et al., 2002; Vargas et al., 2004)

Whole breast ultrasound is sometimes used for the evaluation of suspicious or proven cancer cases as well as for the evaluation of the axilla for possible metastatic nodes. It has also been applied as a screening test for high-risk women (ACRIN 6666 trial) (Berg et al., 2008). The results of the study indicated that incorporating a screening ultrasound with mammography would detect additional 1.1 to 7.2 cancers per 1000 high-risk women, however at the expense of an increased false positive rate. Conventional ultrasound displays the images as thin, flat sections of the body (2D images). Advancements in ultrasound technology include 3D ultrasound that formats the sound wave data into 3D images (Carsten Riis et al., 2005). A recent review of literature on the different kinds of breast ultrasound techniques can be found in Athanasiou et al. (2009). Also, a review of literature on the use of CAD tools on ultrasound images is given in Cheng et al. (2010).

Automated ultrasound: Automated ultrasound is a recent advancement in ultrasonography. Once the technologist positions the transducer over the breast, it automatically performs a preliminary sweeping scan that is used by the technologist to set up imaging parameters for the examined patient. The system then captures many ultrasound images that are rendered and viewed in 3D. Currently, two types of systems are available. One type, the *TechniScan Svara Warm Bath Ultrasound imaging system*, scans the breast that is suspended in a warm water bath held at room temperature. The other type, the *Somo.v system*, features an automated gel coupled scanner. The major advantage of using automated systems is that it reduces the user variability that is widely prevalent in the traditional hand-held methods. Generally, sonographers use automated ultrasound for a good overall view of the breast, and then use traditional hand-held devices to focus on suspicious areas to acquire further details.

Doppler ultrasound: The conventional ultrasound equipment enables Doppler examinations. Doppler ultrasound, in general, is used to detect blood vessels and thereby, to quantify the flow-velocity-based parameters in a tumor. Spectral Doppler and Color Doppler ultrasound utilize the mean frequency shift caused by the Doppler effect of circulating blood cells to examine the vascular flow. The Power Doppler method is based on the measurement of the energy intensity of the Doppler signals, and is effective in examining small lesions (Kook et al., 1999). The results of most of the studies reporting on the use of Doppler ultrasound for breast cancer diagnosis are not consistent. Moreover, studies have reported that Doppler results are greatly affected by the hormonal status, and influenced by the menopausal status and contraceptive usage (Germer et al., 2002). In addition, lesion characterization depends on the proximity of the lesion to the transducer. Other limiting factors are the lack of standardized examination procedures and the prevalent issue of operator dependency of ultrasonography. Hence, currently Doppler imaging is rarely used for breast cancer detection.

Sonoelastography: In sonoelastography, the fact that the breast cancer tissue is about 15 times stiffer than a normal breast tissue is utilized to detect possible cancers. The elasticity of breast tissue is calculated using sonography, and the information is displayed similar to a color Doppler image (Itoh et al., 2006; Scaperrotta et al., 2008). Currently, many clinical ultrasonography systems have the inbuilt capability of measuring the breast tissue elasticity. Sonoelastography is mainly used to improve the accuracy of lesion characterization on ultrasound (Garra et al, 2007).

2.3.5 Computed Tomography (CT)

Computed Tomography (CT, also known as Computed Axial Tomography scan and body section *rontgenography*), is a technique that utilizes X-rays to capture 2D images or slices of the examined anatomy, and applies CAD algorithms to generate corresponding 3D images which provide anatomical information of the positions of air, soft, and bone tissues. In a typical CT facility, the patient lies down on the pedestal and is slowly passed through a ring-shaped apparatus. An X-ray source emits collimated X-rays through the anatomy to be examined, and they are varyingly attenuated or scattered through interaction with the tissues inside. The degree of

scattering depends on the density of the electrons in the tissues. Dense tissues such as bones absorb more X-rays than soft tissues, muscles, adipose tissues, and blood. Thus, a smaller number of X-ray photons reach the detector in the case of bones, rendering a brighter image than soft tissues which appear darker. The detectors on the opposite side record the attenuation as a 1D image. As the X-ray source rotates around the ring-shaped apparatus, these 1D images are shadowed over each other and averaged out to form a single slice 2D X-ray image. After a single full rotation has completed, the patient is moved further according to the desired slice thickness, and another round of slice imaging takes place. In helical or spiral CT scans, the cross-sectional images are taken simultaneously as the X-ray emitting source rotates around the patient in a helical manner continuously while measuring the intensity of transmitted rays from different angles. These segmented slices of images are then collated, stacked upon each other in order to form a volumetric 3D image representation of the internal body organ under scrutiny.

Contrast-enhanced CT: Since CT has a low soft tissue contrast, in order to improve the image quality, certain contrast agents are usually intravenously introduced into the patient. This type of CT is termed contrast-enhanced CT (ce-CT). For breast imaging, intravenous iodinated contrast media are usually used in ce-CT. An example of a contrast enhanced CT scan of the breast is shown in Figure 2.4. The contrast-enhanced CT (axial-acquired sagittal-reconstructed) image shows 8-cm-sized tumor enhancement (arrowheads) in the left breast and enhancing lymph nodes (arrows) without fatty tissue in the left axilla. Contrast-enhanced CT scans are usually performed for staging of metastatic disease and not for staging at the screening or diagnostic level. It was reported recently that a combination of ce-CT and PET whole body scan is efficient in metastatic disease restaging with a lesion detection sensitivity of 93% and specificity of 100% (Dirisamer et al., 2009).

Figure 2.4 Sample contrast enhanced CT image - tumor indicated by arrowheads and enhancing lymph nodes indicated by arrows
(Seo et al., 2005)

2.3.6 Breast MRI

Magnetic Resonance Imaging (MRI) has been approved by the FDA since 1985 as an adjunctive tool, in addition to mammography, to help diagnose breast cancer. Unlike mammography, which uses low dose X-rays to image the breast, MRI uses powerful magnetic fields and radio waves to create images of the breast.

The MRI scanner is a tube surrounded by a giant circular superconductive magnet. During an MRI scan of the breast, the patient lies prone on the scanning table. The breast is inserted into a hole in the table, which contains coils that detect the magnetic signal. The magnet creates a strong magnetic field. When the body is immersed in such a static magnetic field, more protons (which are hydrogen atoms mostly from water present in the body) become aligned with the magnetic field. When a rapidly alternating magnetic field at an appropriate resonant frequency in the Radio Frequency (RF) range is applied, it changes the orientation of the nuclear spins relative to the direction of the static magnetic field. As a result, the nuclei make a transition from a lower energy state to a higher one by absorbing energy from the alternating magnetic field.

When the alternating field is turned off, the nuclei return to the equilibrium state, emitting energy at the same frequency as was previously absorbed. The direction of the main magnetic field is referred to as the Z axis. Before the excitation by RF pulse,

the amplitude in the Z axis is zero and the amplitude in the X-Y plane is maximal. During excitation, the amplitude in the Z axis slowly increases, while the amplitude in the X-Y plane slowly decreases. Therefore, there will be two forms of relaxation. First, the decay of the amplitude in Z axis is known as the T1 relaxation. Second, the re-growth of the amplitude in the X-Y plane is known as the T2 relaxation. Different tissues have different T1 and T2 relaxation rates. Conventional MRI pictures the differences in T1 and T2 relaxation times of tissue water. The MRI image is fairly detailed and hence, tiny changes of structures within the body can be detected. A sample illustration of an MRI image is shown in Figure 2.5.

Figure 2.5 MRI scan of the breast reveals a bright area which proved to be cancer
(Image courtesy of the NCI, VisualsOnline)

Dynamic Contrast Enhanced-MRI: Compared to CT scans, MRI has better soft tissue contrast even without the use of contrast agents. However, due to the heterogeneity of benign and malignant breast lesions, significant overlap was found between the T1 and T2 values of these lesions, thereby necessitating the use of intravenous contrast agents in order to enhance the visibility of the malignant lesions alone (Hendrick, 2007). This enhanced visibility is established by exploiting angiogenesis as an important biological correlate of malignancy. Angiogenesis refers to the formation of new blood vessels and represents a critical step in tumor development. Malignant lesions release angiogenic factors that induce angiogenesis in the surrounding host tissue to establish an independent blood supply that facilitates tumor growth. When a contrast agent is injected, an increased contrast inflow near the tumor enhances the visibility of the

lesion. Most MRI protocols call for the use of T1-enhancing gadolinium-based contrast agents. The contrast agent is chemically bonded in a chelate complex and transported by the bloodstream after intravenous injection. Then, a time series of images are captured. This is the Dynamic Contrast Enhanced-MRI (DCE-MRI) (Heywang-Köbrunner et al., 1997), a technique gaining increased clinical interest because of its ability to provide an enhanced morphology that is the key to differentiating benign from malignant lesions. 3D imaging techniques can also be employed to record a volume of signal data. DCE-MRI has been reported to be more sensitive than mammography (Liu et al., 1998a). An important limitation is the need for specialized software to analyze large volumes of generated data.

A recent study (Lehman et al., 2007) showed that the MRI technique managed to detect cancer in the contralateral breast of 3.1% of women who were diagnosed with cancer in the ipsilateral breast. These women had cancers in the contralateral breast which were missed by preceding mammography and clinical examinations. Such incomplete diagnosis would be health and cost disastrous as the patient might have to endure another round of chemotherapy and other therapeutic treatments.

MRI has been advocated to be utilized as an adjunct to mammographic screening of established high-risk women younger than 40 years of age (Saslow, 2007). This high risk screening utility of MRI is because of its ability to detect early stages of hereditary breast cancers such as intraductal and invasive breast cancers. The diagnosis is achieved at a higher sensitivity as compared with mammography or ultrasound, as the malignant tissues draw a stronger contrast from the surrounding normal tissues in MRI (Warner et al., 2001; Kuhl et al., 2005). Small invasive early-stage tumors in young women are better detected by MRI. Zakhireh et al. (2008) reviewed the literature about MRI for breast cancer screening, diagnosis, and staging.

MR Spectroscopy Imaging: Another important advance in MRI is the Magnetic Resonance Spectroscopy Imaging (MRSI). Numerous studies have demonstrated that malignant tissues have elevated levels of choline-containing compounds, suggesting that these compounds may serve as non-invasive markers for detecting malignancies (Meisamy et al., 2008). *In vivo* non-invasive MRSI uses equipment that is almost identical to the normal MRI apparatus but with specific sequences for spectroscopic

signal acquisition to visualize the total choline content in the breast. Even though the sensitivity of DCE-MRI is high, its specificity varies (30-90%) (Orel et al., 2001). Huang et al. (2004) reported that the addition of MRSI to a normal breast MRI scan improved the specificity from 62.5% to 87.5%. An important limitation of MRSI is its low spatial resolution at traditional field strengths.

CAD in MRI: As the MRI technology continues to advance, the amount of data to be analyzed increases. The number of images to be analyzed may exceed 1000 images per examination. Therefore, the use of computer based image analysis software becomes a necessity to aid in the interpretation of the increasingly complex signals, to reduce the inter-observer variability, and to increase sensitivity and specificity. Typical methods used here are multiplanar reformatting and maximum intensity projection. In general, computerized methods are widely used for CAD and monitoring of treatment. The major components of such methods include reduction of motion artifacts, extraction of temporal features, lesion segmentation, extraction of morphological features, and classification of the features into groups. A detailed review of these methods has been published (Gilhuijs et al., 2008). CAD workstations have been incorporated in DCE-MRI for image subtraction to visualize the enhancement obtained in the images after contrast injection. However, the CAD tools for DCE-MRI are not as well developed as the CAD systems available for mammography. These tools need further development in terms of automated analysis of morphology, multivariate classification, and automated selection of the operating point for CAD-based MRI interpretation.

2.3.7 Positron Emission Tomography (PET)

Positron Emission Tomography (PET) is a cross-sectional imaging technology that allows for the functional imaging of the examined area. Malignant tumors are characterized by increased glucose metabolism. The fluorine in Fluorodeoxyglucose (FDG, 2-fluoro-2-deoxy-D-glucose) molecule serves as the positron-emitting radioactive isotope, fluorine-18, to produce ^{18}F -FDG which is the radioactive tracer commonly used in PET studies. ^{18}F -FDG is taken up by cells with glucose transporters such as GLUT-1, and is converted to FDG-6-phosphate. FDG-6-phosphate can neither undergo further metabolism nor diffuse out of cells. The radioactivity thus

accumulates within the cells. In a low-serum insulin (fasting) state, there is persistent glucose uptake in many cancer cells as opposed to glycogen uptake in normal cells. This produces a good contrast between cancerous and normal cells in PET.

As the accumulated radioactivity decays within the cancer cells, electrons of positive charges, known as positrons, are released. When positrons collide with negatively charged electrons as they move through the tissues, the two particles annihilate each other and release light in the form of two photons, generating gamma rays. These photons move away from each other in opposite directions, and detection of these photons by a ring of detectors in the PET scanning equipment provides information for 3D, cross-sectional representation of the biological activity of the tissues targeted by the radioactive compounds (Moses, 2001). During PET scanning, the patient lies on a table which gradually moves through the PET scanner, typically six to seven times. Clinically, PET is used for locating suspected malignancies and ascertaining the metabolic activities of the detected abnormality and is effective in initial tumor staging and restaging and cancer treatment monitoring. PET has now become an invaluable tool for detecting the recurrence of a tumor based on increased serum tumor marker levels. A sample illustration of a FDG-PET image is shown in Figure 2.6.

Figure 2.6 A PET image showing a focal FDG-avid left breast ovoid mass slightly superior and medial to the nipple (Nofech-Mozes, 2008)

Positron Emission Mammography: PET used in complementary with mammography is known as Positron Emission Mammography (PEM) (Weinberg, 2006). PEM images

are obtained by dedicated breast PET scanners with plate detectors. The detectors are mounted on two compression paddles. As the detectors are placed close to the breast (by gently compressing the breasts), little attenuation occurs, and therefore, the in-plane resolution of PEM is better than that of PET. However, the plate detectors of PEM yield a lower out-of-plane resolution than the circular ring detection arrangement in conventional PET scanners. Also, the gentle compression of the breasts makes registration of PEM images with those obtained with other imaging modalities difficult. A major limitation of PEM is its cost. It also images one breast at a time. In one study, PEM was used to analyze 113 malignancies, and it demonstrated high sensitivity for detecting invasive ductal carcinoma. The results suggested that PEM may not be adversely affected by breast density, hormone replacement therapy, and the menopausal status of the patient (Schilling et al., 2008).

PET-CT: When PET is incorporated with CT, the PET-CT reveals both anatomic and metabolic information in one non-invasive procedure. Combination of PET and CT creates a single or fused image in sequential imaging. CT produces cross-sectional X-rays of anatomical structures, while glucose uptake revealed by PET indicates a metabolically active or malignant tumor. CT alone cannot determine whether a tumor is malignant or benign, but it reveals exactly where the tumor is located. The location information obtained by CT, coupled with the PET information may be of critical value. PET-CT can be a whole body imaging technique to detect metastases spread in the body. PET-CT is now utilized for staging potential metastatic cancers (Avril et al., 2007). A recent study on the applications of PET in breast cancer detection, staging, and therapy assessment can be found in Lavayssière et al. (2009).

2.3.8 Scintimammography

Scintimammography is an adjunctive nuclear-imaging-based breast imaging technique. The patient is administered a single photon-emitting radiotracer. For good tumor differentiation, this tracer has high and specific tumor uptake and minimal activity in the normal breast (Schillaci et al., 2004). The most widely used radiotracers are Tc-99m sestamibi and Tc-99m tetroformin. After the injection of the tracer, the patient lies prone on a special table with the breast hanging down through a hole in the table. The breast is then imaged using a gamma camera at various angles.

The planar method of scintimammography, described above, has low sensitivity for lesions with size < 1 cm (46.5%), and for nonpalpable lesions (61%) (Scopinaro et al., 1997). To improve the sensitivity for these small and non-palpable lesions, tomographic images were acquired with SPECT (Single Photon Emission Computed Tomography) which has better spatial resolution (Schillaci et al., 1997). SPECT is similar to planar scintimammography, but it can provide 3D information by applying tomographic reconstruction algorithms on the acquired multiple 2D images. This technique is best suited for investigating dense breast tissue (Lumachi et al., 2002), breast implants, multiple tumors, and lumps at the site of mastectomy.

Both SPECT and PET are similar in terms of the usage of radioactive tracer material and detection of gamma rays. The basic difference is that in SPECT, the emitted gamma radiation is measured directly, while in PET, the scanner detects the gamma photon emissions that are coincident in time, which provides more radiation event localization information, and thus higher resolution images than SPECT. On the other hand, because of the use of radioisotopes with longer lives, SPECT scans are significantly less expensive than PET scans.

SPECT-CT: Initial clinical investigations using the SPECT-CT system demonstrated that it provides additional value regarding the anatomical location, size and depth of sentinel nodes in invasive breast cancer patients with inconclusive CT findings (Van der Ploeg et al., 2007). Also, the SPECT-CT system has improved sentinel node identification in obese patients (Lerman et al., 2007).

2.3.9 Biopsy

Biopsy is a procedure that removes a sample of the suspicious tissue for further pathological examination. The biopsy is the only way to ascertain whether the tissue sample is cancerous or not. There are several types of biopsies (Joy et al., 2005). In *fine needle aspiration biopsy*, a thin needle is guided into the area of the breast abnormality and suction is applied. Suction will yield fluid or small tissue fragments. Microscopic examination of these samples will determine the nature of the tissue. If the abnormality is too small and not palpable, biopsy can be done with the guidance of conventional imaging approaches. One such procedure, the stereotactic-guided biopsy,

involves the location of the abnormality using an X-ray or ultrasound guided needle to take the sample of a suspicious area. In *core biopsy*, the needle used is larger than that used in fine needle aspiration biopsy, and it removes a small cylinder of tissue from the breast abnormality. *Surgical biopsy* is a technique wherein the entire or a portion of the lump is completely removed using surgical incision. *Axillary node dissection* is a procedure that is used to determine whether breast cancer is present in the lymph nodes. It involves removing 10 to 30 lymph nodes for pathological examination. Another surgical option called *sentinel node biopsy* is now being used on select breast cancer patients to determine whether breast cancer is present in the lymph nodes. A sentinel node biopsy involves removing only one to three sentinel lymph nodes (the first nodes in the lymphatic chain).

2.4 Experimental breast imaging modalities

This section describes in detail the various breast cancer detection modalities that are either in the experimental stage or not being widely used in clinical practice.

2.4.1 Dedicated breast CT

A dedicated breast CT prototype with a high-resolution, isotropic, rotating detector has been developed and tested at the University of California, Davis, USA. During the procedure, a woman lies prone on a table while the imaged breast is inserted into an aperture in the table. The CT scanner then rotates around the breast (Boone et al., 2006). In the initial clinical study, the scanner was used to scan ten healthy women and 69 women with suspicious lesions. The breast CT images were compared with the respective film mammograms. The results suggested that CT offered significantly better visualization of masses, while mammogram was better for microcalcification detection. Without compression of breasts for the CT scan, patients reported better comfort level than mammography (Lindfors et al., 2008). The same research group has also developed a dedicated breast PET-CT scanner and demonstrated its capability of visualizing suspicious lesions (Bowen et al., 2009).

2.4.2 MRI based elastography

In MRI based elastography, a periodic motion (oscillation) is generated by a mechanical shaker to one side of the breast and the resulting displacement field inside

the breast is captured by MRI to determine the elasticity parameters (Van Houten et al., 2003). This technique relies on MRI's ability to detect slight motion.

2.4.3 PET-ER imaging

A novel development in the field of PET is the usage of PET for the *in vivo* quantification of the most widely used therapeutic target, the Estrogen Receptor (ER). ER expression predicts the likelihood of the patient's response to anti-estrogen therapy. [^{18}F]-16 α -[fluororestradiol] (FES) has shown most promise in quantifying the functional ER status of breast cancer, and is thus used as a tracer in PET-ER imaging (Katzenellenbogen et al., 1997). PET-ER imaging (FES-PET) can therefore predict the likelihood of a patient's response to hormonal therapy and thereby to determine the suitability of the patient for this type of treatment. The results of some related studies can be found in Hospers et al. (2008).

2.4.4 Optical imaging

Optical imaging uses light in the Near Infrared (NIR) wavelength zone to scrutinize the inside of the breast, and, thus, is also known as NIR imaging.

Diffuse Optical Imaging: In Diffuse Optical Spectroscopy (DOS), also known as Diffuse Optical Imaging (DOI), the NIR light first penetrates deep into the breast tissues. A series of detectors are used to record the amount of received light and the time taken by the light to pass through the tissues. These recorded measurements are affected by the presence of oxygenated and deoxygenated blood in the studied area. Thus, information related to tissue vascularity and oxygen content can be obtained. Generally, malignancies have greater blood volume and elevated levels of deoxygenated blood relative to benign tissues. Both these markers can be determined by optical imaging. Thus, the DOI technique can detect hyper-vascularized, hypoxic malignant breast tissues.

The NIR light wavelength can be specifically adjusted so that different tissue content can be shown with greater image contrast in preference to surrounding tissues. This variability allows optical imaging to differentiate tumors from the background contrast variations (Quaresima et al., 1998). The DOI technology was implemented in a

commercial system called ComfortScan and distributed by DOBI international, which is now out of business. A systematic review of diffuse optical imaging of the breast can be found in Leff et al. (2008) and in Gibson et al. (2009).

Diffuse Optical Tomography: In Diffuse Optical Tomography (DOT) technique, the target tissue is illuminated by a NIR laser of wavelength in the range of 700 to 1000 nm, and then similar measurements of light signals in the time or frequency domain are made around the tissue boundaries. The lasers encounter low overall tissue absorption, which allows the NIR light to penetrate deeper to allow better deep imaging. These data signals are then compiled and reconstructed to form tissue images. Breast diseases are then identified through comparison and differences in optical properties between the diseased tissues and normal surrounding tissues.

NIR light is highly scattered in the human tissues unlike other imaging techniques such as X-ray, PET, and MRI. Therefore, the signals sometimes become too diffused, and the back-scattering of light might not be detected. Thus, complex numerical methods are used extensively in order to compute the back-scattered light to map out a representation of the inside of the breast tissues. However, recent development in temporal log-slope differentiation mappings has shown greater promise in speeding up this solution for the inverse problem (Guo et al., 2006).

Optical mammography: Optical mammography is a non-invasive technique that uses laser light to illuminate the breast from all sides (Grosenick et al., 2005). Previously, the patient would have been injected with a fluorescent contrast agent. Light absorbed by the contrast agent causes it to fluoresce at a different wavelength, and this emitted light is picked up by highly sensitive detectors around the breast. Using advanced reconstruction algorithms, the collected data is then reconstructed into a 3D image of the breast tissue. The main advantage of this technique over the traditional mammography technique is that no ionizing harmful X-ray radiation is used here. The optical mammogram (mediolateral projection) of a patient with multi-centric invasive ductal carcinoma on the left breast is depicted in Figure 2.7.

Figure 2.7 Optical mammogram of the left breast in mediolateral projection; the patient has multi-centric invasive ductal carcinoma as indicated by the arrows (Grosenick et al., 2005)

2.4.5 Thermography

Thermography is based on the principle that the temperature of the area surrounding the pre-cancerous and cancerous breast tissue is higher than that of the area around the normal breast tissue. Cancerous and pre-cancerous tissues are highly metabolic, and hence, they usually have a constant growth of new blood vessels to supply nutrients to the rapidly growing cancer cells. Due to this angiogenesis, the surface temperature of these tissues tends to be higher than usual. Breast thermography uses a highly sensitive, high-speed infrared camera to capture a digital image of the infrared heat radiated off from the breast. Human skin emits infrared radiation mainly in the wavelength range of 2 to 20 μm and with an average peak at 9 to 10 μm (Archer et al., 1971). The imager converts the invisible thermal energy to electrical signals in order to display the temperature profile of the subject. The temperature profile is an illustration of many colours each of which indicates a particular temperature. A computer-assisted interpretation of the digital image then helps to determine whether a local abnormality in breast tissue temperature is present or not.

Thermography is passive in nature as it will not emit any harmful radiation or subject the patient to any risk. In addition, it is a non-contact screening process, and is useful for real-time imaging. The thermography systems have better portability compared to many other modalities. However, there are numerous concerns when it comes to clinical implementation and practice. The thermal imager should be operated in a

stable indoor environment with stability of the operating ambient temperature (18-25°C) within $\pm 1^\circ\text{C}$, and with facilities to control all ambient conditions such as temperature, humidity and additional infrared sources. Relative humidity is to be kept from 40 to 75% (Ng, 2009). The accuracy of the thermal imager is highly dependent on the skill and knowledge of the operator (Ng et al., 2001; Standards Technical Reference, 2003). Also, different scanners have different operating parameters, thereby necessitating the need for good knowledge about the device used. Thus, even though thermal imagers offer an excellent means of making a qualitative determination of surface temperature, there are many difficulties in obtaining an absolute measurement.

Some researchers have concluded that the thermogram provides a reflection of functional tumor induced angiogenesis and metabolic activity (Keyserlingk et al., 2006), and hence, the information from thermograms could be considered as a significant biological risk marker for cancer (Keith et al., 2001; Amalu et al., 2006). A detailed review of thermography and its efficiency and limitations can be found in Ng (2009).

2.4.6 Electrical characteristics based imaging techniques

Over the years, several *in vitro* and *in vivo* studies have been conducted to study the electrical characteristics of cancerous and non-cancerous breast tissues. These studies indicate that malignant tissues typically have higher conductivity, and/or permittivity, or lower impedivity when compared to the normal and benign tissues. In this section, first, an overview of the electrical impedance and electropotential characteristics of normal and malignant breast tissues is presented. Second, the efficacy of a few devices based on these characteristics, such as T-Scan 2000, T-Scan 2000ED, MEIK, and BDS, is assessed using the clinical trials carried out to date. Since this technique is the key basis for this work, a detailed review is given.

2.4.6.1 Electrical impedance characteristics of breast tissue

Tissues consist of cells and extracellular medium. The cell consists of the cell membrane and the intracellular medium. Both the extracellular and intracellular media

contain ions and they are highly conductive in nature. The cell membrane, which is made up of lipid bilayer and proteins, is primarily capacitive. At low frequencies, the impedance associated with this capacitance becomes so high that the cell membrane starts behaving like an insulator. Hence, the current flows only through the extracellular space and the resulting impedance is largely resistive. At higher frequencies, the cell membrane starts conducting, and hence, the impedance is lower. Thus, it is evident that tissues exhibit frequency dependent behaviour (Holder, 2005). When the relative permittivity of a biological tissue is plotted against frequency f it forms a piecewise linear function. The plot is depicted in Figure 2.8.

Figure 2.8 Ideal representation of permittivity as a function of frequency
(Pethig, 1987)

As seen from the figure, there are three distinguishable regions called the relaxation regions (Foster et al., 1989; Pethig, 1987; Schwan, 1957; Schwan et al., 1993). They are called alpha ($f < 1000$ Hz), beta ($1000 \text{ Hz} < f < 100 \text{ MHz}$), and gamma ($f > 1 \text{ GHz}$) regions. The alpha region depicts the conductance values of the ionic extracellular fluid. The beta region depicts the permittivity values that arise due to the increasing capacitive charging and discharging of the cell membranes and the dipolar relaxation of proteins. The gamma region represents the permittivity due to bipolar relaxation of the water. This frequency dependent behaviour of the tissues necessitates the need to record the impedance over several frequencies. These observations lead to the conclusion that different components of the tissue contribute to different impedance values, which also vary with different frequencies. Hence, in any impedance based application, it will be logical to obtain measurements over several frequencies. There is also a significant difference between the impedance of normal

living tissue and that of dead tissue. This is because the permeability of the cell membrane changes within hours of cell death (Polk et al., 1986). Hence, it is better, if the impedance measurements are made *in vivo* or within minutes of death, to obtain accurate information about the tissue.

To understand the impedance characteristics of normal/benign and malignant breast tissues, many *in vitro* and *in vivo* studies have been carried out over the past few years. These studies have investigated the impedance changes over a wide frequency range, covering all the three relaxation regions. Following is the summary of the results, categorized by frequency range.

Results of impedance studies done at low frequencies (< 100 MHz)

Fricke and Morse (1926) were the first to report the significantly higher permittivity of the tumor tissue at 20 kHz compared to that of the normal or benign tissues. Singh et al. (1979) performed *in vivo* impedance measurements on female breasts with and without tumors in the frequency range of 100 Hz - 100 kHz. The results showed that malignant tumors had higher relative permittivity and lower resistance than those of normal breast tissue. In 1984, Chaudhary et al. examined excised normal and malignant breast tissues in the frequency range of 3 MHz - 3 GHz, and concluded that the conductivity and permittivity of malignant tissues are higher than those of normal tissues, especially at frequencies less than 100 MHz.

Surowiec et al. (1988) carried out *in vitro* dielectric studies in three different samples of tissues, the main tumor tissue, the tissue immediately surrounding the tumor, and the peripheral healthy tissue at the frequency range of 20 kHz - 100 MHz. The results revealed the existence of significant difference in the dielectric properties among the various tissue samples. The tumor tissues recorded a conductivity of around 2 to 3 mS/cm at 100 kHz, which is higher than the conductivity of the normal tissue (1 mS/cm). Cellular heterogeneity and structural differences among the tested samples were stated as the reason for this difference.

Morimoto et al. (1990, 1993) measured the extracellular and intracellular resistances and membrane capacitance of normal tissues and breast tumors *in vivo* over a frequency range of 0 - 200 kHz. Statistically significant differences between normal

and cancerous tissues were reported. Moreover, the malignant tumors were reported to have lower capacitance compared to that of benign tumors, contrary to the findings of Fricke et al. (1926) and Jossinet (1996) that the malignant tumors had higher capacitance values.

An *in vitro* study of six groups of normal and pathological tissues at 12 discrete frequencies halving from 1 MHz to 488 Hz was carried out by Jossinet (1996, 1998, 1999). In one study, the variability of impedance values within each group was assessed by statistical methods (Jossinet, 1996). It was found that the variability was smaller in adipose tissue, carcinoma, and fibroadenoma at frequencies above 10 kHz. In another study (Jossinet, 1998), using the same data, it was found that the cancerous tissue differed significantly from fibroadenoma and mammary gland tissues by the modulus of impedivity up to 31.25 kHz, and from the remaining tissue groups (connective tissue, adipose tissue, and mastopathy) by the low-frequency limit resistivity and the phase angle from 125 kHz to 1 MHz. In the third study, Jossinet et al. (1999) defined a set of eight parameters that could differentiate cancerous tissues from non-cancerous ones. It was concluded that a combination of the parameters over various frequencies is necessary for the accurate differentiation among tissues. Chauveau et al. (1999) conducted an *in vitro* study of normal and pathological breast tissues at frequencies ranging from 10 kHz to 10 MHz. Three indices were determined based on extracellular resistance, intracellular resistance, and membrane capacitance to differentiate various tissue pathologies such as normal, Invasive Ductal Carcinoma (IDC), IDC with stromal reaction, and fibrocystic changes.

Results of impedance studies done at high frequencies (> 100 MHz)

Campbell et al. (1992) conducted an *in vitro* study of the dielectric properties of fatty, normal, and cancerous breast tissue samples at microwave frequency (3.2 GHz). The dielectric properties of all the different tissue types were reported to be similar contrary to the findings recorded in other such studies listed earlier. The overshadowing effect of polar water on the changes that result from the biologically different heterogeneous structures may be the reason for the observed results (Rigaud et al., 1996).

Joines et al. (1994) measured the properties of normal and malignant tissue types of the colon, kidney, liver, lung, and mammary gland at frequencies from 50 MHz to 900 MHz. A significant increase in the permittivity and conductivity of malignant tissues was observed when compared to that of the normal ones. Among all the tissue types, the largest difference was recorded for the mammary tissue.

In 2007, Lazebnik et al. (2007a) measured the dielectric properties of 354 excised normal breast tissues over the 0.5 - 20 GHz frequency range. The tissues were categorized into three groups based on the percentage of the adipose tissue content. The dielectric constant and effective conductivity of the normal tissue samples were reported to increase with the decrease in the adipose content. In another study by Lazebnik et al. (2007b), the microwave dielectric properties of a variety of normal, malignant, and benign breast tissues were investigated at the same 0.5 to 20 GHz frequency band. The dielectric properties of benign and normal breast tissues with low adipose content were found to be similar. Moreover, a high contrast in the dielectric properties between malignant and adipose-dominant normal tissues (in the ratio of 10:1) was reported, which is in agreement with the results of Joines et al. (1994). However, the contrast in the dielectric-properties between malignant and normal glandular or fibroconnective tissues (low adipose content) was much less (consistent with the results of Campbell et al. (1992)). These two observations indicate that the large variability in the dielectric properties of normal tissue resulting from tissue heterogeneity should be considered while developing breast cancer detection devices working in the microwave frequency range.

It has been inferred from the results of the above mentioned studies that the malignant breast tissue typically has higher conductivity, and/or permittivity, or lower impedivity when compared to normal or benign tissues especially at frequencies lower than a few MHz. This frequency range has also been validated by Blad and Baldetorp (1996), who suggested that alpha and beta dispersion regions are particularly useful for tumor detection measurements since most of the differences between normal and pathological tissue seem to appear in this frequency range. In another study, Blad et al. (1999) determined an electrical impedance index, which is sensitive to the variation in characteristic frequency of the tissues, to differentiate cancerous tissues from normal tissues. However, a common index for all the cancer types has been proposed. It has

been suggested that with the characteristic frequencies of each individual cancer type known, their respective indices can be fine-tuned for accurate classification.

Pethig (1984) attempted to analyze the characteristics of malignant tissues in order to determine the characteristics that were the cause for the observed difference in the dielectric properties between normal and malignant breast tissues. Some of the characteristics stated include higher water content and sodium concentration of cancerous cells, and different electrochemical properties of their cell membranes. Other studies (Foster et al., 1989; Morimoto et al., 1993; Jossinet, 1996) indicated that apart from the increased cellular water and salt content, the altered membrane permeability, changed packing density, and orientation of cells of malignant tissues were the reasons for the difference in dielectric properties.

2.4.6.2 Electropotential characteristics of breast tissue

Normal/Benign tissue: Normal breast epithelial cells are electrically polarized by an ionic gradient across the cell membrane (Davies, 1996). The charge gradient is asymmetric between the apical (inside lining of the cell membrane) and basolateral (at the base) cell membranes, and is known as the Transepithelial/ transmembrane Electrical Potential (TEP) (Goller et al., 1986). The charge gradient across the membrane is maintained by the different permeabilities to ions, predominantly K^+ and Na^+ ions, and this gradient causes a TEP difference of about 30 mV (Figure 2.9). As epithelial cells divide, this charge gradient across the epithelial layer is dissipated due to the movement of ions thereby resulting in depolarization – (1).

Malignant tissue: When a cancer develops, the epithelial cells in cancerous areas divide more rapidly than those in the other areas of the breast resulting in a pocket of relative depolarization. As the cancer progresses, the basement membrane of cell is invaded by more malignant cells and the ductal structure inside the breast is destroyed. The permeability of the malignantly altered cells to Na^+ is increased, and Na^+ enters the cell increasing the intracellular Na^+ concentration while the extracellular K^+ concentration increases. Thus, there is more movement of these ions in the areas of malignancy – (2)

Figure 2.9 Breast epithelial electrical gradients
(Pethig, 1987)

Thus, in both cases (1) and (2), there is movement of ions. Movement of ions equals electric current. The transmembrane electropotentials generated by these ionic currents have generally been measured by using microelectrodes inserted through the cell membrane, and they aid in the understanding of biopotentials and the nature of the cell processes and functions and correlate with surface measurements, such as the electrocardiogram (ECG). However, skin surface electropotential measurement techniques do not measure the transmembrane potential. They measure the signal that is generated externally from cells due to the external flow of current from a region of the cell membrane having a higher voltage, to one having a lower voltage; the signal is thus due to the current flow through the conducting medium. If such a signal exists in the steady state, a DC signal will be generated.

To detect this DC signal, a biopotential sensor is placed on the skin. This sensor has an electrolyte (gel) and an electrode (made of silver-silver chloride layers) on top of the gel. When the gel comes in contact with the skin, the ions from the skin surface are driven into the gel by electromotive force because of the tendency of the ions to move from higher to lower concentration to attain equilibrium. These gel ions then attract the silver chloride ions in the electrode. The gel ions and the discharged silver chloride ions form an electrical double layer. Layer 1 contains ions in the gel and layer

2 which is parallel to layer 1 contains ions in the silver chloride electrode. This is similar to a capacitor with a little distance in between the two ion layers. Thus an electric field is created between these two layers resulting in the measured electropotential. Thus, even without applying any external current/voltage stimulation, the naturally occurring inherent biological electricity is recorded by the sensors (passive sensing). Apparently, the BDS sensing system (explained later) uses digital filters at appropriate frequencies to filter out the ECG AC signals that might come along with the DC electropotentials when the sensors are used in the chest area.

Even if the cancerous lesion is situated deep inside the breast and is not palpable, it is still possible to differentiate it, because cancer cells exhibit the property of field carcinogenesis (Gallager et al., 1969). It means that the abnormal cell growth is manifested even at a distance further away from the actual tumor leading to altered surface electropotentials over the entire field. These altered surface potentials can be indicative of pre-cancerous conditions, thereby warning women about probable cancer risk (Cuzick et al., 1998).

2.4.6.3 Electrical impedance scanning

Electrical Impedance Scanning (EIS) is a non-invasive *in vivo* imaging technique used to map the electrical conductivity and capacitance values of tissues for the purpose of differentiating malignant and normal tissues. In EIS, a constant AC voltage is applied to the body between a scanning probe containing an array of sensing electrodes and a reference electrode. The currents through the electrodes are then measured. With the current and potential known at each electrode, the impedance of each electrode is then calculated. These values are then projected to the surface and mapped as 2D EIS images.

T-Scan 2000

The only FDA-approved EIS device is the TransScan T-Scan 2000, manufactured by Mirabel Medical Systems. The pre-market approval from FDA was received in April 1999 (http://www.accessdata.fda.gov/cdrh_docs/pdf/P970033b.pdf). This device has been approved for use as an adjunct to mammography in patients who have equivocal mammographic findings within BI-RADS category 3 or 4 (BI-RADS explained at a later section). A positive T-Scan finding in the vicinity of a suspicious lesion

strengthens the necessity for biopsy, while a negative supports other management options. The T-Scan 2000 system, shown in Figure 2.10, is a real-time, non-invasive, radiation-free, 2D, multi-frequency imaging device for breast examination that maps the local distribution of tissue electrical impedance.

Data acquisition

In a T-Scan 2000 examination, an AC voltage of amplitude between 1 and 2.5 volts with frequencies spanning between 100 Hz and 100 kHz is applied between the hand-held reference electrode and the scanning probe placed over the breast. Although a large range of frequencies are used during the test, the interpretation is made only from the data obtained using FDA-approved 200 Hz (Malich et al., 2003a). The examination is generally performed with the women recumbent, with the arm on the same side of the examined breast raised above the head. This position will flatten the breast as much as possible, allowing a good contact between the flat surface of the measuring probe and the underlying breast tissue. To obtain a better contact between the probe and the breast surface, ultrasound gel is used. The measuring probe contains a multitude of electrodes arranged on a rectangular grid and each of the electrodes is held at virtual ground. The evoked currents are measured at the electrode pads. The calculated capacitance and conductance values are then mapped to gray scale values - higher values are mapped to lighter shades of gray and the lower values to darker shades - and then displayed on the monitor. A detailed description of the operational principle of T-Scan is given in Assenheimer et al. (2001).

Figure 2.10 T-Scan 2000

The T-Scan 2000 system comes with three different probes (Assenheimer et al., 2001). The small probe has an eight by eight pad array with a total footprint of 32 mm by 32 mm. The large probe consists of 16 by 16 pad elements with a footprint of about 72 mm by 72 mm. The latest probe has the same footprint as the small probe, but has 16 by 16 pad elements. The area occupied by all the sensor elements is optimized for generation of homogeneous electric field under the probe. T-Scan 2000 can be operated in two modes. The *targeted mode* (Figure 2.11) is used only for examining the location of an already known suspicious mammographic or sonographic lesion. In this mode, each of the sensors in the probe individually measures the impedance values at the targeted location. The *scanning mode* or the *standard resolution mode* (Figure 2.12) is used for acquiring the impedance maps of both the entire breasts. The probe is moved in a pre-defined sequence over the entire breast. In both the modes, two three by three data sectors are recorded - one sector

represents the capacitance map and the other the conductivity map. The center of each map represents the values obtained from the nipple.

Figure 2.11 Screen display in the targeted mode of T-Scan 2000

Figure 2.12 Screen display in the scanning mode of T-Scan 2000

The imaging is real-time and the probe can be moved over the breast to obtain the best possible image. The processing software allows the operator to adjust the gray-scale mapping levels to highlight certain features alone. Because breast cancer cells have a conductance higher than that of the normal breast cells, breast tumors may appear as bright white spots on the computer screen. According to cell properties, conductivity

of nipple tissue is high compared to that of the rest of the breast tissue, and hence, a healthy nipple is represented as a bright white spot. Thus, when a normal breast is imaged using T-Scan, the only area that should appear as a bright spot is the nipple. If there are other bright spots in the image sectors, they may be due to a clinical finding, artifact due to skin lesions, or abnormal. T-Scan can detect cancerous lesions as small as 3 mm in diameter (Malich et al., 2003b). A typical breast examination performed with the T-Scan 2000 system is shown in Figure 2.13.

Figure 2.13 A typical breast examination using T-Scan 2000
(Zou et al., 2003)

Clinical trials

The robustness and limitations of the T-Scan 2000 device and the algorithm used can be evaluated using the data obtained from clinical trials carried out over the past few years. Table 2.1 gives a brief summary of most of the trials and the respective results.

Table 2.1 Summary of clinical trials carried out using T-Scan 2000

Authors (Year)	Nature of study	Study results and comments
Piperno et al. (1990)	Studied the suitability of impedance measurements for breast cancer screening using a prototype to the T-Scan 2000 called Mammoscan.	No statistical results like sensitivity or specificity were reported.
TransScan group (http://www.accessdata.fda.gov/cdrh_docs/pdf/P970033b.pdf)	Carried out three studies to evaluate the effectiveness of T-Scan 2000 as an adjunct to mammography.	In one blinded study, the sensitivity of 60% obtained using mammography alone increased to 82% on using adjunctive T-Scan (combination of T-Scan and mammography). The specificity increased from 41% to 57%. In another study using the targeted mode, the sensitivity and specificity were around 78% and 67%, respectively. A higher sensitivity of 93.3% and specificity of 83.3% were achieved when targeted mode of T-Scan 2000 was used and results interpreted adjunctive to mammography results.
Scholz et al. (2000)	Conducted numerical simulations to determine the parameters that affect the impedance measurements.	The measured peak signal of a lesion of a fixed size at a fixed distance from the measurement plate decreased with increasing breast thickness. In the case of large breast, the electric field became inhomogeneous and the lesion signals kept decreasing with increasing surface area and edge effects also started appearing.
Pertlet et al. (2000)	Investigated the correlation between hormonal factors and the false positive results associated with EIS using the T-Scan 2000 in 21 healthy pre-menopausal women.	Each woman was examined once per week for two menstrual cycles. The number of white spots that appeared in the T-Scan image varied with the menstrual cycle and had temporal instability. They concluded that the presence of these false positive spots did not strongly correlate with age, hormonal use (oral contraceptive use) or the electrical characteristics of the breast. Similar research in large varied population is necessary to conclude if these results can help to distinguish true positive and false positive spots on T-Scan images.
Malich et al. (2000)	Conducted a study to evaluate the efficiency of T-Scan 2000 on sonographically/ mammographically suspicious lesions.	A sensitivity of 93.1% and specificity of 65.5% was achieved for the targeted mode. The standard resolution mode had lowered sensitivity of over 75% and specificity of 72.4%.

Literature review

Authors (Year)	Nature of study	Study results and comments
Malich et al. (2000a)	Evaluated the advantage of combining EIS, mammography and ultrasound.	The addition of EIS to mammography and ultrasound increased the sensitivity from 86.4% to 95.1% while decreasing the accuracy from 82.3% to 75.7%. They concluded that EIS is a promising technology since it increases the sensitivity when combined with mamumography and/or ultrasound; but its clinical use depends on the reduction of false positive results.
Malich et al. (2000b)	Evaluated 100 mammographically suspicious lesions using EIS, ultrasound, Magnetic Resonance Mammography (MRM) and biopsy.	With EIS, a sensitivity and specificity of 81% and 63% were observed. Since the results of EIS were found to be comparable to MRM results, the suggestion was that EIS can be used as an adjunctive technology in situations where MRM cannot be used.
Martin et al. (2002)	Carried out T-Scan 2000 examination on mammographically suspicious lesions to understand the relation between the histopathology of abnormal breast conditions and intensity, depth and multiplicity of EIS signals.	Multiple EIS signals were observed more frequently in cases with malignant biopsy than in benign biopsy cases. The EIS signal intensity did not decrease with the increasing depth of the lesions (in the depth range 0.5 - 5.5 cm). This is contradictory to the results published by Scholz et al. (2000). This difference may be attributed to the geometrical breast model errors and simplicity of the breast conductivity distribution assumed in the numerical model. This group also observed that benign proliferating lesions were a reason for false positive results.
Piperno et al. (2002)	Conducted a clinical trial to evaluate the use of electrical impedance measurements to determine the estrogen activity in 86 postmenopausal women.	The conductance and capacitance values decreased with the increasing number of years from the onset of menopause. It was noticed that nipple conductance measured at 200 Hz was higher in women taking Estrogen Replacement Therapy (ERT) than those who weren't taking ERT, thereby indicating that EIS can be a suitable method for predicting estrogen activity.
Malich et al. (2003b)	Studied the influence of size and depth on the accuracy of targeted mode of T-Scan 2000.	Large homogeneously structured lesions did not present a visible focal spot because of the increase in the conductance of the measuring field, and therefore, the decrease in the relative changes in the larger field. In such cases it was observed that, if the border of the lesions were examined, the sensitivity obtained will be comparable to that obtained using other lesion sizes. They concluded that the software should have options for depth adjusted impedance determination and should also take into account the size parameter.

Literature review

Authors (Year)	Nature of study	Study results and comments
Malich et al. (2003a)	Conducted the EIS study on 240 women with 280 lesions and presented a review on the clinical value and limitations of EIS.	<p>An overall sensitivity of 80.5% and specificity of 64.7% were recorded. They observed the following</p> <ul style="list-style-type: none">• When the examination frequency was increased to higher values the bright spots switched into black areas.• A large number of malignant lesions did not register a focal increased value of capacitance; the reason for this has to be investigated.• The absolute values of the conductivity and capacitance recorded largely depend on the pressure between the breast and the scanner and critically depend on the operator's skill and experience.
Zhenyu et al. (2005)	Attempted to study the parameters influencing EIS imaging using simulations.	<p>When the tumor was close to the measuring electrodes, the electrical disturbance signal was large and the resulting image was clear and vice versa. Keeping the distance of the tumor from the measuring electrodes constant, it was found that with increasing tumor size the contrast of the resulting image was lost. Poor contact between the measurement probe and skin surface resulted in artificial imaging.</p>
Glickman et al. (2002)	In order to stabilize the variations in the sensitivity and specificity values recorded in the above studies, this group developed an objective nonuser-dependent post processing algorithm to standardize the data interpretation process.	<p>Thresholds were calculated using the crossover frequency and phase of the electrical impedance parameters recorded over the entire frequency range (100 Hz -5 kHz) at only the lesion and nipple region and not the entire breast. The results are displayed as Level of Suspicion (LOS) scores ranging from 1 to 5 for each recorded image, indicating the likelihood of a malignant finding in the image. This algorithm was verified on a biopsy identified test set of 87 cancerous and 153 non-cancerous cases and 356 screening cases. 84% sensitivity and 52% specificity were observed</p>
Szabo et al. (2005)	Conducted the EIS exam (targeted mode combined with the post-processing LOS software described earlier (Glickman et al., 2002) on 145 lesions found by palpation or mammography.	<p>EIS recorded a sensitivity of 86% and a low specificity of 49%. They suggested that proliferative benign disease or hormonal status of patients could be the reason for the lowered specificity. This group concluded that the targeted T-Scan could not provide significant adjunctive diagnostic aid to mammography and ultrasound because of its similar sensitivity and lowered specificity when compared to those of mammography and ultrasound.</p>

Literature review

Authors (Year)	Nature of study	Study results and comments
Diebold et al. (2005)	Performed an EIS exam on 256 preoperative patients with BI-RADS III/IV/V-lesions using the targeted mode of T-Scan 2000 and the LOS software.	It was found that EIS had a sensitivity of 75.4% and a specificity of 42%. Sensitivity changed with lesion size - 85% for size lesser than 10 mm and 64% for size greater than 10 mm. The low values of sensitivity and specificity were attributed to the fact that a large percentage of the lesions in their study were located at a greater depth, not easily identifiable by EIS. They concluded that the LOS software setup provided better improvement over the usual T-Scan focal spot based method, especially for small malignant lesions.
Fuchsjaeger et al. (2005)	Attempted to study the negative predictive value of T-Scan 2000 (using the LOS software) in analyzing 128 BI-RADS category 4 breast lesions.	A high negative predictive value of 97.1% was obtained. This indicates that instead of subjecting the BI-RADS 4 lesions which had negative EIS result to biopsies, the physician can treat these lesions as those belonging to BI-RADS 3 with confidence and ask the patient to come for a short-term follow-up.

T-Scan 2000ED

T-Scan 2000ED is yet another device from the developers of T-Scan 2000. It received the CE mark in 2003. It is primarily a risk assessment tool intended to be used as a complement to CBE in asymptomatic women who are between 30 and 39 years of age with a negative CBE result and negative family history of breast cancer.

Data acquisition

T-Scan 2000ED uses the same measurement technique, patient interface, and electrical current levels of T-Scan 2000, but uses a different scoring algorithm that maximizes specificity so that it can be suitable as a screening tool. Dickhaut et al. (1996) suggested that the measured impedance values reflect more widespread electrical changes within the breast thereby indicating precancerous changes even in the absence of a lesion. This finding strengthens the suitability of EIS for use in screening younger women who normally do not have palpable lesions. Hence, this new device uses the data obtained from the entire breast and processes them in a manner to maximize specificity and minimize the number of false positives. Since electrical impedance tends to decrease with age, the decision threshold for this device was set at a higher level. This implies that the sensitivity of T-Scan 2000ED would be lower for older women. More details on the algorithm are given in (Stojadinovic et al., 2008).

In T-Scan 2000ED, the conductance and capacitance values are recorded over a wide frequency range of 100 Hz to 1 GHz. T-Scan 2000ED displays only the conductance map. This map is only used for proper positioning of the scanning probe and is not for clinical interpretation. The conductance and capacitance values measured at all frequencies are processed to give the final result. The result is presented on a binary indicator bar (green indicates negative and red indicates positive). A positive result indicates that the woman is at a higher cancer risk and helps the physicians decide on further breast examination tests that are required.

Clinical trials

In 2005, Stojadinovic et al. (2005), attempted to evaluate the efficiency of T-Scan 2000ED in 1103 women. Sensitivity and specificity, for the sub group of women younger than age 40, using this new device were 50% and 90%, respectively. In 2006,

Stojadinovic et al. conducted a two-arm, prospective, multi-site clinical trial to evaluate the use of EIS to identify younger women at high risk of cancer. The sensitivity arm had 189 women aged 30-45 years with suspicious breast abnormality who underwent CBE and EIS (T-Scan 2000ED) prior to biopsy. A sensitivity of 38% was recorded in this arm. In the specificity arm, a total of 1361 women aged 30-39 years were enrolled for the study. A high specificity of 95% was recorded as was expected in a T-Scan 2000ED examination. These results were confirmed by another prospective, two-cohort trial conducted by Stojadinovic et al. in 2008. It was found that T-Scan 2000ED was able to identify a subset of young women with a relative probability of breast cancer almost five times greater than in the population of young women at large. Since the device is not intended to be used for lesion localization, if a patient is found at a higher risk, she should undergo other diagnostic procedures for further evaluation. Thus, T-Scan 2000ED is intended to be used as a screening tool to select at-risk young women for highly sensitive breast cancer diagnostic tests. Even though T-Scan 2000 and T-Scan 2000ED have undergone many clinical studies and have obtained government approvals for commercial distribution and usage, both the devices are not in the market presently.

2.4.6.4 Electrical impedance tomography

In EIT, a large number of impedance measurements are made from electrodes placed on the body surface, and the results are processed by a computer to produce reconstructed tomographic 2D or 3D images of the impedance. To obtain an image with good spatial resolution, a number of such measurements are obtained by applying different current distributions to the body, and repeating the voltage measurements. These set of measurements are then used to generate the tomographic image by complex image reconstruction techniques. EIS technique is almost similar to EIT, except in the electrode arrangement and reconstruction algorithms. In most EIT systems, the electrodes are arranged such that they surround the entire object that is being imaged, for example, the electrode array will be in a circular fashion for imaging the breast. However, in EIS, the electrode arrangement is planar, thereby avoiding the necessity of complex reconstruction algorithms that are usually needed for EIT.

In EIT, it is possible to inject a range of currents at different frequencies (Bayford, 2006). A variety of experimental EIT systems have been developed worldwide. Detailed information on EIT is presented in Bayford (2006) and Holder (2005). PKF "SIM-technika" Ltd., a company in Russia, developed a commercial EIT device called Mammograph Electrical Impedance Komputer (MEIK) (Figure 2.14), the details of which are presented in this section.

Figure 2.14 Mammograph Electrical Impedance Komputer
(Image courtesy: http://www.impedance.ru/en/product_opis.html)

Data acquisition

The CE-approved MEIK system provides a 3D view of the electrical conductivity distribution in the breast tissues. For scanning purpose, the device utilizes a constant current of 0.5 mA at 50 kHz. Initially, two disposable reference electrodes are attached on the inside of the wrist of the hand nearest to the breast to be examined. During examination, the scanning probe, which consists of 256 electrodes arranged in a matrix, is pressed against the breast so as to increase the number of electrodes in contact with the breast and to decrease the thickness of the tissue layer to be measured. Once the scanning process is initiated by pressing a button, each of the 256 electrodes, one at a time, successively injects current into the examined area and the corresponding surface electropotentials are recorded. This entire scanning process takes about 35 seconds.

The 3D distribution of the electrical conductivity is then constructed by using the method of weighted back projections along the equipotential surfaces of the electrical field. Seven tomographic planes are reconstructed within a time span of 20 seconds (Figure 2.15). The resulting image represents tissue conductivity along a grey scale

from dark to bright, which represents low to high conductivity. Tumors and nipple areas, being highly conductive, appear as bright areas. The processing software has many inbuilt capabilities such as numeric data filtering, color scaling, contrasting, highlighting of a specific area, computation of mean electroconductivity, minimum and maximum values of the electroconductivity, standard deviation and modes, plotting of electrical conductivity distribution, and comparative analysis with normal database. There is also an option for comparing the electroconductivity of the left and right breasts using the values of the percentage of spread in the electroconductivity distribution.

Figure 2.15 Data presentation screen in MEIK
(Image courtesy: http://www.impedance.ru/en/product_opis.html)

Clinical trials

Currently, the data provided by the manufacturer indicate that the device has a sensitivity of 92% and specificity of 99%. Very few clinical study groups have published their results. Table 2.2 summarizes the results of the few clinical trials.

Table 2.2 Summary of clinical trials carried out using MEIK

Authors (Year)	Nature of study	Study results and comments
Cherepenin et al. (2001)	Evaluated an earlier version of MEIK, termed as Breast Cancer Detection Device (BCDD) (0.2 mA, 10 kHz) on 21 women with tumors in one breast. The device used 256 electrodes arranged in a square matrix and used the method of weighted back projection for image reconstruction.	The images called Electrical Impedance Mammograms (EIM) were reconstructed every 8 mm down to a depth of 6 cm. They found that 86% of the diagnoses made by BCDD were fully or partially consistent with diagnoses made by other tests. They have attributed the data mismatch to enlarged breasts which might affect the results due to limited scanning depth and also to the effects of medical treatments such as puncture, biopsy, or x-ray therapy done before the BCDD test in certain cases. They found a significant difference in electrical conductivity between the normal and tumor affected breasts at all scanning planes. In the regions estimated solely to correspond to tumor, increased conductivity was observed at all scanning planes.
Cherepenin et al. (2002)	Evaluated the possibilities of using BCDD (0.5 mA, 50 kHz) in the examination of mammary glands with different hormone status. Instead of the square matrix, a round matrix was used to allow more electrode contact with the skin and to increase image resolution.	They found that the EIT images were the same at the first and second phases of the menstrual cycle and also during pregnancy. During lactation, they observed light colors at the center of the image reflecting the increased conductivity due to the milk content in the ducts. The EIT images at menopause were distinctly different from the other images with the increased presence of lighter areas. They also observed similar conductivity levels at the first and second phase of a 28-day menstrual cycle and increased conductivity in postmenopausal breasts. The presence of lighter areas in the EIM of postmenopausal breasts and their increased conductivity can be attributed to their increased fat and connective tissue content. They concluded that all these observations about the healthy breasts at various physiological periods can serve as a benchmark when pathological breast tissues are evaluated at a later stage.
Korotkova et al. (2007)	Described the procedure that is most suitable for interpreting the EIM images obtained from MEIK.	They enlisted various visual details and quantitative parameters that might be valuable in accessing the images.
Sotskova et al. (2007)	Attempted to determine the particularities of the EIM images in different forms of infiltrative breast cancer.	The EIM of complicated forms of breast cancer revealed deformation of the mammary gland contour, subcutaneous fat infiltration, perifocal infiltration, and anatomic changes with displacement of internal structures. Non-complicated nodular forms of breast cancer presented no visual changes in their EIMs.

2.4.6.5 Biofield Diagnostic System (BDS)

The Biofield Diagnostic System (BDS), developed by Biofield Corporation, USA (<http://www.mackaylifesciences.com/products.html>), provides an immediate, objective assessment of the nature of the breast tissue *in vivo* with the help of measured skin surface DC electropotentials. This device received market approval from the European Union (CE mark) for use as an adjunct to physical breast examination or relevant imaging modalities to detect breast cancer in patients less than 56 years of age with palpable lesions.

It works on the principle that the irregular rapid cell division associated with cancer causes the area around the cancer to be more depolarized than the nearby normal tissue, thereby resulting in large electropotential differences between the cancerous and normal areas of the breast (Davies, 1996; Goller, 1986; Marino et al., 1994). These differentials that are transmitted to the breast surface by means of volume conduction can be measured using specially designed sensors and measurement device (Weiss et al., 1994; Faupel et al., 1995; Sacchini, 1996; Faupel et al., 1997; Faupel et al., 1998; Cuzick et al., 1998; Cuzick et al., 2002). Also, because of the concept of field carcinogenesis (Gallager et al., 1969), which states that the abnormal cell growth is manifested even at a distance further away from the actual tumor, the altered electropotentials arising out from deeply situated malignant areas could be measured at the skin surface. A detailed description of the principle and electropotential measurements was given in section 2.4.6.2. The BDS device uses eight single-use sensors for each breast and two reference sensors to measure these skin surface electropotentials. It was observed that cancers produced significantly greater electropotential differentials than benign lesions (Sacchini, 1996; Faupel et al., 1997; Cuzick et al., 2002).

Data acquisition

The BDS device is a desktop sized electronic device with a display screen, keyboard, printer, and patient cables for each breast with attached sensor leads (Figure 2.16). During the examination, each of the nine leads on each cable is connected to nine single-use sensing electrodes (pads) that are pre-gelled. Each biopotential sensing electrode consists of a sensor body that is formed from a round sensor disc and a mounting pin which projects outwardly from the sensor disc. A slice of plastic foam is

attached to the lower surface of the disc. The plastic foam is impregnated with an electroconductive medium (in the form of a gel) that encounters the skin surface. Each of the sensor leads is clamped to the mounting pin during measurement. The diameter of the sensor disc is between 0.39 and 0.43 inches. The sensor body is formed from a non-metallic material (glass-plastic combination) and is uniformly coated with a metallic coating (silver-silver chloride layer).

Figure 2.16 The BDS system

On each of the leads, there is a visual code to assist in proper positioning relative to the location of the lesion. A lead clamped to one sensing electrode pad makes up one sensor. The positioning is done as follows: one diagnostic sensor is placed over the center of the suspicious lesion on the symptomatic breast (SC) and four additional sensors are placed outside the margins of the lesion - superior to (SU), inferior to (SL), medial to (SI), and lateral to (SO) the sensor placed over the test lesion. Two additional sensors (SH and SV) are placed in the center of the two quadrants (horizontal and vertical directions) adjacent to the symptomatic quadrant and one sensor is placed in the quadrant diagonal (SD) to the quadrant with the test lesion. This pattern of sensor placement is then replicated in a mirror image form on the asymptomatic breast, whose sensors are labeled as AC, AU, AL, AI, AO, AV, AH, and AD accordingly. This type of placement (as depicted in Figure 2.17) is to make sure that the opposite breast serves as a control for evaluating the asymmetry between

the two breasts. Two reference sensors are placed just below and between the breasts. Once the patient demographic parameters like age and menstrual history are keyed in, the test begins and goes on for around seven minutes in order to allow the device to take repeat measurements of the electropotentials and finally average them to reduce the noise level. After the test is completed, the sensor pads are removed and the results are printed.

Figure 2.17 Sample illustration of the sensor arrangement used on the patient for the BDS test, assuming that the suspicious lesion is in the upper outer quadrant of the right breast

The sensors are accurate, and low in noise and electrical impedance. The device uses microprocessor technology to manipulate the signals acquired - sampling, filtering, and data processing. Once the sensor impedance values are acquired, the device calculates and outputs a BDS index between +5 and -5 (This is different from the T-Scan and MEIK devices, which display an image). This ***BDS index*** is combined with the ***prior Level Of Suspicion (LOS)*** values assigned by the physician after clinical diagnosis (ultrasound, mammography etc), using an interpretation map described later, to determine an adjusted LOS, called ***Post-BDS LOS***, which will provide guidance for further diagnostic procedures. The main requirement of BDS is that it should be used on a patient with a suspicious lesion that has been located by palpation or by using

other diagnostic procedures. Thus, BDS is intended to be used as an adjunctive diagnostic test.

Variations in the BDS device

The BDS system has undergone a few improvements in its interpretation algorithm and type of sensor arrangement since its inception in the early 1990s. For ease of usage, the device used for CE studies (BDS Physician manual) will be henceforth referred to as the *CE device*, and that used for Lancet studies (Cuzick et al., 1998) will be called *Lancet device*.

BDS index calculation: In earlier versions of both Lancet and CE devices, the BDS index was calculated by the system using the following variables - SC (value from the central sensor which is placed over the suspicious lesion), SMED (median of the values from sensors in the uninvolved quadrants of the symptomatic breast), ARANGE (range of the values from sensors in the asymptomatic breast), the age of the patient, and the palpability status of the lesion. The variables SC, SMED, ARANGE, and age were treated as continuous variables. The palpability variable was treated as binary (0=non-palpable, 1=palpable). The dependent variable was a binary outcome (malignant or benign) as determined either by biopsy or by clinical evaluation. Logistic regression was then used to model the logit transformation of the probability of malignancy as a function of these variables - SC, SMED, ARANGE, Age, Palpable and Age*Palpable. The resulting BDS index, which is the logit transformation of the probability of malignancy, was displayed after each test (BDS Physician manual). Let the above algorithm be denoted as *Algo 1*. In the latest versions, like the device used in this work, the Palpable and Age*Palpable parameters were not included (Sacchini et al., 2002). Let this modified algorithm be denoted as *Algo 2*.

BDS interpretation map: Earlier Lancet and CE devices used a Clinical Management System (CMS) (Sacchini et al., 2002) for interpretation. The CMS was devised with input from clinical scientists active in the field of breast oncology. The input to the CMS system is the BDS index obtained after conducting the test. The following are the CMS thresholds used to define the three CMS categories:

- Green: $\text{BDS index} < -1.36$: *Favorable*
- Yellow: $-1.36 \leq \text{BDS index} \leq -1.17$: *Borderline*
- Red: $\text{BDS index} > -1.17$: *Unfavorable*

This categorical output permits integration with the BI-RADS scales commonly used for mammography results. Table 2.3 depicts the CMS based *Map 1* used in clinical practice to determine the LOS of the patient (BDS Physician manual). The clinician first determines the appropriate prior LOS based on other imaging results (mammography or ultrasound) and other clinical information for that patient. The BDS index or CMS color obtained from the patient’s Biofield test is then located in the first column. The intersection point of the prior LOS column and the CMS row indicates the revised LOS called Post-BDS LOS. However, the device used in this work uses modified thresholds, depicted in the *Map 2* presented in Table 2.4 (BDS Physician manual). The prior LOS values 4 and 5, however, should not be adjusted based on the BDS index. This is because these lesions have already been declared to be highly probable of being cancerous with great confidence and hence, any lower Post-BDS LOS should not alter the prognosis and inflict patient risk.

The prior Level of Suspicion (LOS) definitions for breast lesions follow generally accepted guidelines:

- **LOS 1:** Negative/ Normal/ No Finding – The breasts are symmetrical and no masses, architectural distortions, or suspicious calcifications are present.
- **LOS 2:** Benign – Imaging or physical examination findings may be present, but there is no suggestion of malignancy.
- **LOS 3:** Probably benign/ Malignancy not excluded – Findings suggest a high probability of being benign, but malignancy cannot be ruled out. The lesion is not expected to change over the follow-up interval.
- **LOS 4:** Probably Malignant/ Suspicious Abnormality – These lesions do not have the characteristic morphologies of breast cancer but have a definite probability of being malignant.
- **LOS 5:** Malignant/ Highly Suggestive of Malignancy – These lesions have a high probability of being cancer.

These LOS definitions are similar to the definitions used in BI-RADS. BI-RADS is an acronym for Breast Imaging Reporting and Data System. It refers to the mammography assessment categories used by a radiologist while interpreting a mammogram. The BI-RADS assessment categories are: 0- Incomplete, 1 - Negative, 2 - Benign finding(s), 3 - Probably benign, 4 - Suspicious abnormality, 5 - Highly suggestive of malignancy, 6 - Known biopsy, proven malignancy.

Table 2.3 BDS interpretation map 1: using CMS thresholds

Clinician should determine the prior LOS below					
	LOS 1	LOS 2	LOS 3	LOS 4	LOS 5
	Normal/No Finding	Benign	Malignancy Not Excluded	Probably Malignant	Malignant
Biofield CMS					
FAVORABLE (green) BDS index < -1.36		Benign	Malignancy Not Excluded	Probably Malignant	Malignant
BORDERLINE (yellow) -1.36 ≤BDS index ≤ -1.17		Benign	Probably Malignant	Malignant	Malignant
UNFAVORABLE (red) BDS index > -1.17		Probably Malignant	Probably Malignant	Malignant	Malignant

Table 2.4 BDS interpretation map 2: using modified thresholds

	LOS : 1	LOS : 2	LOS : 3
	Normal/ No finding	Probably benign	Malignancy not excluded
BDS index			
-5.0 to -1.7	Probably benign	Probably benign	Probably benign
-1.6 to +0.2	Probably benign	Probably benign	Malignancy not excluded
+0.3 to +1.5	Malignancy not excluded	Probably benign	Probably malignant
+1.6 to +5.0	Malignancy not excluded	Malignancy not excluded	Probably malignant

Sensor arrangement: Also, the CE and Lancet devices varied in terms of placement of the reference electrodes and the equilibration period prior to data acquisition. In the Lancet device, one reference sensor was placed on each palm, seven sensors on each breast, and one on each axilla (*SA 1*). In the CE device, two reference sensors were

placed in the subxiphoid area and eight sensors on each breast (SA 2). The differences and consequences are highlighted in Table 2.5. As observed, the CE device produces lesser noise, and has a quicker data acquisition period. Moreover, it is unaffected by electropotentials from the sweat glands. Clinical trials demonstrated that a lower equilibration period was sufficient and could provide results similar to that obtained using the ten-minute period (Faupel et al., 1997).

Table 2.5 Differences between CE and Lancet BDS devices

CE DEVICE	LANCET DEVICE
Subxiphoid (below the xiphoid, which is the posterior and smallest of the three divisions of the sternum) reference; used to reduce noise	Palm reference
Diagonal sensor in lieu of axilla sensor	Axilla sensor; reason for increased electropotential values possibly due to sweat glands
Five-minute equilibration period prior to data acquisition	Ten-minute equilibration period prior to data acquisition
Age group tested in the CE study : 18-55	Age group tested in the Lancet study : 18-90

Clinical trials

BDS is still not available for commercial distribution in the U.S. because it is yet to receive FDA’s pre-market approval. Since it has not yet been approved by FDA, there have not been as many clinical trials as those undergone by T-Scan. BDS received market approval in the European Union (CE mark) after conducting a clinical trial on a total of 463 patients falling within the age bracket of 18-55 years. There were 299 patients who underwent biopsies of which 150 were benign and 149 malignant. A comprehensive algorithm implemented in this study using the Biofield test yielded a sensitivity of 98.7% and specificity of 52.2% (BDS Physician manual). It was also demonstrated that BDS increases the malignancy detection by up to 20-29% compared to the current diagnostic techniques such as mammography and ultrasound.

The output of the CMS index when combined with the mammography LOS using Table 2.3 gives a modified LOS classification (CMS + LOS), also denoted as the Post-BDS LOS. In a pivotal study conducted by BDS (BDS Physician’s manual) for the purpose of results submission to FDA, subgroup analyses were done to assess the

effect of age, lesion palpability and size, and nature of cancer (invasive or non-invasive) on the CMS+LOS results. It was found that CMS+LOS provided added cancer detection (increase of 22% or more) for all age groups compared to mammography LOS alone. Hence, the adjunctive use of the BDS for all age groups is positive and supportive. CMS+ LOS detected 33.2% (79/238) more palpable cancers than mammography alone. The results showed that higher predictive information occurred in women with palpable lesions. In the case of lesion size, cancer detection for CMS+LOS remained high (90% or higher) for all ranges of lesion sizes, the smallest being less than 0.5 cm. The sensitivity with CMS+LOS versus mammography LOS alone increased for both non-invasive (40% higher) and invasive lesions (23% higher). All the above results prove that CMS+LOS combination methodology provided higher effectiveness within these difficult sub-populations, and hence, the adjunctive effectiveness of BDS to mammography is promising and supportive.

The Post Market Surveillance (PMS) work (BDS Physician's manual) was carried out as a routine practice after BDS obtained the CE marking. 389 women, who satisfied the inclusion and exclusion criteria, participated in the test. PMS results of 90% sensitivity and 40% specificity demonstrate the utility of the Biofield test under prospective utilization in the clinical setting for breast cancer detection.

In a clinical trial conducted by Cuzick et al. (1998), 661 patients above the age of 18 years (18 to 90 years), who had a suspicious lesion scheduled for excision biopsy, underwent the BDS test which was conducted at eight centers in five European countries. It was concluded that BDS provided highly significant discriminatory information for women with palpable lesions. This group has developed a depolarization index (BDS score) based on the electropotential differences and age. The median depolarization index was 0.40 for nonproliferative benign lesions, 0.53 for proliferative benign, and 0.64 for cancer. Using age and this index, a specificity of 55% was obtained at 90% sensitivity for women under 56 years with palpable lesions. In the case of non-palpable lesions, only a minimal discrimination was achieved. The reason for this is attributed to the difficulty in accurate sensor placement over the non-palpable lesions.

Ki-67 antigen expression is a commonly used invasive method for distinguishing aggressive cancers (with high proliferative activity) and malignant lesions that have proliferative activity similar to benign lesions (which appear less than 10% of the times). To confirm that the BDS test measures net electrical activity related to cellular proliferation as determined by this Ki-67 antigen expression, a prospective study of 103 consecutive patients with palpable breast lesions and scheduled for open biopsy was conducted (Gatzemeier et al., 2002). In this study, the average Ki-67 antigen expression values were compared with preoperative BDS index using Spearman and Wilcoxon rank correlation. The results showed a positive correlation with the Ki-67 antigen expression. Dickhaut et al. (1996) reported that the BDS score correlated with the Thymidine labeling index, another marker for breast cancer, in 84 patients with cancers at a test conducted in Istituto Nazionale Tumori in Milan, Italy. In another study in Japan, 101 women were tested using BDS and sensitivity and specificity values of 90% and 60%, respectively, were obtained (Fukuda et al., 1996). Thus, all the above clinical trials demonstrated the effectiveness of BDS as an adjunct modality to other diagnostic techniques such as mammography and ultrasound.

There have been studies using BDS in a few other countries in the recent years. The results of these studies have not yet been published in peer reviewed journals. A summary of the various BDS preclinical and clinical studies, along with the features of the BDS device used in each study, is given in Table 2.6.

Literature review

Table 2.6 Summary of the BDS clinical trials

Year of study	Source	Sample size	Sensor arrangement	Age range	Lesion palpability	Interpretation	Results
1994	BDS report (Sacchini, 1996)	104	SA 1 ^	≥ 18 years	Palpable and non-palpable	Prospective algorithm based on electropotentials	Sensitivity: 97% Specificity: 79%
1995	BDS Lancet study (Cuzick et al., 1998)	661	SA 1 ^	≥ 18 years	Palpable and non-palpable	Discriminant based on depolarization index and age	AUC : 0.8440 (for palpable lesions)
1996	Fukuda et al., 1996	101	SA 1 ^	≥ 18 years	Palpable and non-palpable	Threshold found using a probabilistic neural network built on electropotential differentials	AUC: 0.783
1995-1998	BDS CE study (BDS Physician manual)	463	SA 2 ^^	18-55 years	Palpable	Using BDS index calculated using <i>Algo 1*</i> and prior LOS on <i>Map 1</i>	AUC: 0.915
1995-1998	BDS Pivotal study (BDS Physician manual)	617	SA 2 ^^	≥ 18 years	Palpable and non-palpable	Using BDS index calculated using <i>Algo 2**</i> and prior LOS on <i>Map 1</i>	AUC: 0.8987
1995-1998	BDS Pivotal subset study (BDS Physician manual)	396	SA 2 ^^	18-55 years	Palpable and non-palpable	Using BDS index calculated using <i>Algo 2**</i> and prior LOS on <i>Map 1</i>	AUC: 0.8724
1995-1998	BDS PMS study (BDS Physician manual)	389	SA 2 ^^	18-55 years	Palpable and non-palpable	Using BDS index calculated using <i>Algo 2**</i> and prior LOS on <i>Map 1</i>	AUC: 0.77
2009	Clinical study in this work	182	SA 2 ^^	≥ 18 years	Palpable and non-palpable	Using BDS index calculated using <i>Algo 2**</i> and prior LOS on <i>Map 2</i>	AUC: 0.972

**Algo 1* - Logit transformation of the probability of malignancy as a function of values recorded by the sensor over the suspicious lesion (SC), median of the values recorded from sensors in the uninvolved quadrants of the symptomatic breast (SMED), the range of values from sensors in the asymptomatic breast (ARANGE), age and lesion palpability status (BDS Physician manual)

***Algo 2* - Linear combination of SC, SMED, ARANGE and the age of the patient (Sacchini et al., 2002)

^ SA 1: One reference sensors on each palm, seven sensors on each breast and one on each axilla

^^ SA 2: Two reference sensors in the subxiphoid area and eight sensors on each breast

2.4.7 Other experimental modalities

Microwave imaging spectroscopy (Fear et al., 2002) is an examination which involves the propagation of low levels of microwave energy through breast tissue to measure the electrical properties. Generally, tumors have been found to have more water and blood than the regular tissue. Since this technique is particularly sensitive to water, it is able to distinguish malignant and benign tumors. Ultrasound CT is another emerging modality that captures two sets of images, one based on the speed of sound and the other based on the attenuation of sound. Each set is made up of a collection of coronal slices of the breast. A summary of a few other alternative breast imaging modalities can be found in Brenner et al. (2007). A review on the new modalities for breast cancer detection can be found in Singh et al. (2008).

2.5 Comparison of the various breast imaging modalities

A comparison of the clinical and experimental breast imaging modalities, in terms of resolution, advantages, limitations, and key uses in breast cancer imaging, is presented in Table 2.7.

Table 2.7 Comparison of the various breast imaging modalities

Imaging modality: Results of well-controlled clinical trials, if available	Spatial resolution	Advantages	Limitations	Key uses in breast imaging
Film Mammography: 66% sensitivity and 92% specificity (Pisano et al., 2005).	Superior spatial resolution - around 20 line pairs per mm (20 lp/mm) $\approx$ pixels 25 μ m apart (Hashimoto, 2008).	<ul style="list-style-type: none">• Most widely accepted imaging modality with good cancer detection accuracy.• Standard for breast cancer screening.• Minimal examination time.• Effective in detecting microcalcifications	<ul style="list-style-type: none">• Use of ionizing radiation.• Reduced cancer detection accuracy in dense breasts relative to fatty breasts.• A physical film archive is required.• No form of image post-processing can be done.• Requires repositioning of the breast in the mammogram machine for each view.• Some women experience discomfort during breast compression.	Used for breast cancer screening and diagnostic purposes.
Digital Mammography: 70% sensitivity and 92% specificity (Pisano et al., 2005).	Lower spatial resolution compared to its film counterpart - around 5-10 lp/mm $\approx$ pixels 50 - 100 μ m apart (Hashimoto, 2008) and higher contrast resolution.	<ul style="list-style-type: none">• Most widely accepted imaging modality with good cancer detection accuracy.• Standard for breast cancer screening.• Reduction in physical film storage space.• Efficient and safe handling of the images.• Image enhancement to accentuate pathology is possible.• Utilization of CAD software to highlight suspicious areas possible.• Quick image transmission to get rapid second opinions.• Minimal examination time.• Effective in detecting microcalcifications.	<ul style="list-style-type: none">• Use of ionizing radiation.• Reduced cancer detection accuracy in dense breasts relative to fatty breasts.• The device is about two to three times more expensive than a film mammography machine.• The doctors who were previously used to reading film-screen mammography have to undergo training to read the digital mammography images and trained to use the mammography workstation.• Requires repositioning of the breast in the mammogram machine for each view.• Some women experience discomfort during breast compression.	<ul style="list-style-type: none">• Used for breast cancer screening and diagnostic purposes.• Used for breast biopsy guidance.

Literature review

Imaging modality: Results of well-controlled clinical trials, if available	Spatial resolution	Advantages	Limitations	Key uses in breast imaging
Breast ultrasound: 97% accuracy (Lister et al., 1998).	Relatively lower spatial resolution compared to mammography (Lagalla et al., 1998).	<ul style="list-style-type: none">• Relatively less expensive.• Uses no ionizing radiation.• Real-time imaging.• No breast compression.• Provides a clear picture of soft tissues.	<ul style="list-style-type: none">• Microcalcifications are not easily visualized.• There may be image degradation in technically difficult patients, for example obese patients.	<ul style="list-style-type: none">• Used for typical cancer diagnostic purposes and breast biopsy guidance.• Preferred to mammography in the evaluation of young patients with dense breasts.
CT Scan	High spatial resolution but not as high as that of mammography.	<ul style="list-style-type: none">• Relatively shorter scanning time compared to other modalities.• Excellent for seeing bones, vessels, and organs and the precise location and size of a mass.	<ul style="list-style-type: none">• Use of ionizing radiation.• Inability to differentiate between a benign or malignant mass, scar tissue or cancerous tissue.	<ul style="list-style-type: none">• Not used routinely for the assessment of breast nodules.• Occasionally used in the assessment of the primary breast tumour in cases of suspected advanced local invasion.• Used for staging and restaging of cancer.
SPECT: 92.9% sensitivity and 87.5% specificity (Schillaci et al., 1997).	Lower resolution than PET.	<ul style="list-style-type: none">• Because of the use of radioisotopes with longer lives, SPECT scans are significantly less expensive than PET scans.	<ul style="list-style-type: none">• Lower resolution than PET.	Best suited to investigate dense breast tissue, breast implants, multiple tumors and lumps at the site of mastectomy.

Literature review

Imaging modality: Results of well-controlled clinical trials, if available	Spatial resolution	Advantages	Limitations	Key uses in breast imaging
MRI: 88.1% sensitivity and 67.7% specificity (Bluemke et al., 2004).	High resolution	<ul style="list-style-type: none">• Ability to acquire images in any plane and any orientation without having to reposition the patient.• No ionizing radiation.• High sensitivity.• Well suited to image dense breasts (Park et al., 2006).• No breast compression.• Angiogenesis detection using DCE-MRI results in high sensitivity.• Excellent sensitivity and specificity presented by choline-MRSI.	<ul style="list-style-type: none">• Inability to detect microcalcifications• Long acquisition time.• Requires patients to remain still during the scan as images are susceptible to motion artifacts.• Scanning of patients with pacemakers, cochlear implants and certain other metallic foreign bodies is contraindicated due to the presence of high magnetic fields.• Expensive.• Less specific, resulting in many false positives indicating the need for unnecessary biopsies.	<ul style="list-style-type: none">• Used to evaluate the integrity of breast implants.• Used in pre-operative planning as it can help surgeons determine the extent of the primary breast tumor and if other suspicious foci are present in a better way.• Used for post-surgical evaluation.• MRI guided breast biopsy for lesions visible on MR but occult on MG and breast ultrasound.
Optical Imaging	Low resolution	<ul style="list-style-type: none">• Relatively less expensive.• Uses no ionizing radiation.• Provides cellular and molecular information.	<ul style="list-style-type: none">• NIR light is highly scattered, and hence, complex numerical methods are required for image reconstruction.	Adjunct tool for the detection and characterization of breast lesions.
Thermography: Average sensitivity and specificity of 90% (Amalu, 2006).	Low resolution	<ul style="list-style-type: none">• Relatively less expensive.• Uses no ionizing radiation.• No breast compression.	<ul style="list-style-type: none">• Accuracy is highly dependant upon the operator and interpretations are difficult.• Stable indoor environment with ambient temperature and humidity conditions are vital for good accuracy.	Adjunct tool for the detection and characterization of breast lesions.

Literature review

Imaging modality: Results of well-controlled clinical trials, if available	Spatial resolution	Advantages	Limitations	Key uses in breast imaging
PET: 48% sensitivity and 97% specificity (Kumar et al., 2006).	<ul style="list-style-type: none">• Whole body PET: low resolution of about 6 mm (Tartar et al., 2008).• PEM: higher resolution of about 2 mm (Tartar et al., 2008).• Relatively poorer resolution compared to CT and MRI.	<ul style="list-style-type: none">• No breast compression.• Effective in detecting metabolically active tumors.• Can be used in conjunction with CT for a whole body scan in detecting metastases.• Molecular probes for tumor characterization based on ER are under development.	<ul style="list-style-type: none">• Time consuming.• False positive results may occur.• Not effective in picking up microscopic metastatic foci.• Expensive.• A cyclotron facility is needed in/ near the hospital for producing the short- lived radioisotopes.	<ul style="list-style-type: none">• Useful for staging. Distant metastases can be detected.• Used for re-staging of patients with a local/regional recurrence or metastases.• More accurate in detecting larger and more aggressive tumors than in locating tumors that are smaller than 8 mm and/or less aggressive.• PET is able to provide information on the metabolic activity of the tumour and hence, is helpful in monitoring treatment effectiveness.

Literature review

Imaging modality: Results of well-controlled clinical trials, if available	Spatial resolution	Advantages	Limitations	Key uses in breast imaging
Electrical Impedance Tomography (EIT)	Low resolution	<ul style="list-style-type: none">• Uses no ionizing radiation.• Good temporal resolution.• No breast compression.	<ul style="list-style-type: none">• Varying electrode-skin contact impedance.• Limited number of electrodes.• Poor signal to noise ratio.• Complex reconstruction algorithms.	Adjunct tool for the detection and characterization of breast lesions.
Electrical Impedance Scanning (EIS) T-Scan 2000: 80.5% sensitivity, 64.7% specificity (Malich et al., 2003a) T-Scan 2000ED: 50% sensitivity and 90% specificity (Stojadinovic et al., 2008)	Low resolution	<ul style="list-style-type: none">• Uses no ionizing radiation.• Capability to image dense tissues.• More objective interpretation.• Less expensive.• Fast examination time.• No breast compression.	<ul style="list-style-type: none">• False positives are caused by unequal gel thickness between probe and breast, artifacts such as bones, muscles, and skin lesions.• Cancerous lesions directly behind the nipple are not always detectable due to the similarity in appearance on the EIS image.• Not possible to localize positive lesion for biopsy.	Adjunct tool for the detection and characterization of breast lesions.
Biofield Diagnostic System (BDS)	Low resolution	<ul style="list-style-type: none">• Uses no ionizing radiation.• Capability to study dense tissues.• More objective interpretation.• Less expensive.• Fast examination time.• No breast compression.	<ul style="list-style-type: none">• The accuracy of the sensor placement is dependent on the operator as there are no images to guide the operator for accurate placement.• The reasons for false positives have to be investigated for better accuracy of the device.• More clinical trials are needed to evaluate the device.	Adjunct tool for the detection and characterization of breast lesions.

2.6 Statistical analysis techniques and definitions

This section presents the definitions and descriptions of the various statistical parameters used in this work.

2.6.1 Performance measures: contingency table

Several parameters can be used to describe the quality and usefulness of a diagnostic test. Among those, the most commonly used parameters for clinical trial results analysis are sensitivity, specificity, and accuracy. *Sensitivity* is the probability that a test will produce a positive result when used on diseased population. *Specificity* is the probability that a test will produce a negative result when used on disease-free population. Accuracy is the ratio of the number of correctly classified samples to the total number of samples. In the context of the BDS test, the actual disease status is the final diagnosis predicted for the patient by the doctor using other tests or biopsy. The BDS test result is considered as negative if the Post-BDS LOS value is less than or equal to 3, and positive if it is greater than 3. I.e., Post-BDS LOS 3 is used as the cut-off point or *decision threshold*. The formulae to calculate these parameters are shown in the contingency Table 2.8.

Table 2.8 Definition of performance measures

	ACTUAL DISEASE STATUS	
TEST RESULT	Malignant	Benign
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)
PERFORMANCE MEASURES		
Sensitivity	TP/(TP+FN)	
Specificity	TN/(TN+FP)	
Accuracy	(TP+TN)/(TP+TN+FP+FN)	

- True Positive (TP) - Number of diseased patients for whom the test results were positive
- True Negative (TN) - Number of disease-free patients for whom the test results were negative
- False Positive (FP) - Number of disease-free patients for whom the test results were positive
- False Negative (FN) - Number of diseased patients for whom the test results were negative

2.6.2 Performance measures: AUC

Another important performance measure is the *Area under the Receiver Operating Characteristic (ROC) curve*, called AUC. The ROC curve is obtained by calculating the sensitivity and specificity of a diagnostic test at different cut-off values and

plotting sensitivity vs. (1-specificity) (Metz, 1978). (1-specificity) is called the False Positive Rate (FPR). A test that perfectly discriminates between the two groups (benign and malignant) would yield a curve that coincides with the left and top sides of the plot. This means that sensitivity is high and the FPR is low. A test that is completely useless would give a straight line that follows a diagonal path from the bottom left corner to the top right corner. This means that for every improvement in sensitivity there is an equal increase in false positive rate or decrease in specificity. Generally, the curve will lie somewhere between these extremes because of the overlap of the values in the two groups. A good diagnostic test is one that has small false positive and false negative rates across a reasonable range of cut-off values.

Generally, the goodness of a diagnostic test is assessed by determining the Area under the ROC curve (AUC). For an ideal test, the AUC would be 1, meaning sensitivity and specificity of 1 (100%). For a useless test, which follows the diagonal ROC curve, the AUC would be 0.5 which is equivalent to having sensitivity and specificity of 0.5 (50%). Hence, in practice, the closer the AUC is to 1.0, the better the test is, and the closer the AUC is to 0.5, the worse the test is (Zweig et al., 1993). An AUC value of 0.9 indicates that a randomly selected individual from the positive group has a test value larger than that of a randomly chosen individual from the negative group 90% of the time.

While plotting the ROC curve for diagnostic tests that are based on BI-RADS scale, like the BDS test, one can choose from four different cut-off levels to define a positive test for malignancy. They are ≥ 2 (i.e., the most liberal criterion), ≥ 3 , ≥ 4 , and ≥ 5 (i.e., the most stringent criterion). Therefore, there are four pairs of sensitivity and specificity values, one pair for each cut-off level. These pairs can be placed on a graph with the sensitivities as the y-coordinates and the (1-specificities) as the x-coordinates. The ROC curve can then be constructed by joining all the four points. The resulting ROC curve is called the empirical ROC curve (Figure 2.18). Because the ROC curve displays the sensitivities and FPRs at all possible cut-off levels, it can be used to assess the performance of a test independent of the *decision threshold* (Park et al., 2004).

Figure 2.18 Sample ROC curve

2.6.3 Correlation coefficient

The correlation coefficient indicates the strength and direction of a linear relationship between two variables. The correlation is closer to 1 in the case of an increasing linear relationship and closer to -1 in the case of a decreasing linear relationship. The closer the coefficient is to either -1 or 1 , the stronger the correlation between the variables. A value of 0 indicated the absence of a linear correlation. It is important to note that a significant correlation between two variables does not imply a causal relationship between them. Also, a non-significant correlation coefficient only indicates that there is no linear relationship between the variables; the variables may have a non-linear relationship.

Two types of correlation coefficients are widely used in practice. The *Pearson's correlation coefficient* (r) is a parametric correlation coefficient that is used to study the relationship between two continuous variables that are normally distributed. *Spearman's correlation coefficient* (ρ) is a rank correlation coefficient that is used on two ordinal variables or on two continuous variables that do not have a normal distribution. In this method, the variables are first ranked, and the ranks are used to calculate the correlation.

2.6.4 t-test

A two-sample t-test is a parametric test used to estimate whether the mean value of a normally distributed outcome variable (result) is significantly different between two groups of participants. A one-tailed t-test uses the null hypothesis that mean_1 is greater than mean_2, i.e., it is used to test the effect in one direction only. A two-tailed t-test, on the other hand, uses a null hypothesis that states that there is no difference between the two means. Once a t-value is determined, a p-value is found using a table of values from the t-distribution. Then the determined p-value is compared with a level of significance (α -level). Popular levels of significance are 5% (0.05), 1% (0.01) and 0.1% (0.001). If the p-value is lower than the α -level, the null hypothesis is rejected. This means that the means are not equal (in the case of two-tailed test), and mean_1 is lesser than mean_2 (in the case of one-tailed test).

2.7 Data mining techniques and definitions

This section presents the definitions of the techniques used in the various data mining steps such as data pre-processing, feature selection, classifier development, and classifier evaluation. A dataset of size $n \times m$ has n rows and m columns. In medical datasets such as the datasets used in this work, each row represents one patient record; each column refers to one feature like age. Thus, values are recorded for all features of all patients in a dataset.

2.7.1 Data pre-processing: Handling of missing values

Feature values that have not been measured in some participants are called missing values. To maintain the study power while reducing any bias in the summary statistics, it is a common technique to replace a missing value with an estimated value. This can be done by replacing the missing value by the mean value or the median value.

2.7.2 Data pre-processing: Outlier detection

Outliers are data values that are extreme compared to the other values in the data set. It is important to detect outliers because they can change the result of the data analysis. There are two types of outliers: univariate outliers and multivariate outliers. A univariate outlier is a data point that is quite different from the rest of the data for one feature. An outlier is measured by the distance from the remainder of the data in units

of the standard deviation, which is the standardized measure of the spread of the data. In this work, the univariate outliers were detected using z-score method. To calculate the z-score for an observation under an attribute, the mean of the attribute is subtracted from the observation, and then divided by the standard deviation of the attribute. If the absolute value of the z-score is greater than 3, then the observation is a univariate outlier. If the sample size is small (80 or fewer cases), the z-score threshold is kept at ± 2.5 .

A multivariate outlier is a case that has an unusual combination of values for a number of features. The value for any one of the individual features may not be a univariate outlier, but, in combination with the values from other features, it might become an outlier. Mahalanobis D^2 is a multidimensional version of a z-score. It measures the distance of an observation from the centroid (multidimensional mean) of a distribution, given the covariance (multidimensional variance) of the distribution. A case is a multivariate outlier if the probability associated with its D^2 is 0.001 or less. Most authors using this distance for outlier detection propose a p-value of 0.001 (Peat et al., 2005; Foster et al., 2005). D^2 follows a chi-square distribution with degrees of freedom equal to the number of variables included in the calculation.

2.7.3 Data pre-processing: Normality tests

Statistical methods are based on various underlying assumptions. One common assumption is that a random variable is normally distributed. A normal distribution is classically a bell shaped curve that is bilaterally symmetrical. If a variable is normally distributed, then the mean and the median values will be approximately equal. There are two broad techniques used for testing whether a continuous variable is normally distributed. They are numerical tests and normality plots. Normality plots used are histograms, normal Q-Q plot, detrended normal Q-Q plot, and box plot. Numerical tests use statistical parameters like skewness and kurtosis, or *Tests of Normality* like Kolmogorov-Smirnov statistic and Shapiro-Wilk statistic. The Shapiro-Wilk statistic is based on the correlation between the data and the corresponding normal scores. It has been found that this test has more statistical power to determine a non-normal distribution than the Kolmogorov-Smirnov test (Stevens, 1996). Hence, this test was used in this work. Generally, the presence of outliers causes the attribute to be non-

normal. Hence, after the removal of outliers, the normality of the attribute should be checked for improvement.

2.7.4 Data pre-processing: Normalization

Data normalization, in this context, refers to the scaling of data into a particular range. This step is important when attributes (also called as features) are used as inputs for classifiers. The same range of values for each input feature is preferred to minimize bias for one feature over another within the neural network. In this work, min-max normalization technique was employed to re-scale each attribute to have a range between 0 and 1. Min-max normalization performs a linear transformation on the original data. Let $\min A$ and $\max A$ be the minimum and maximum values for attribute A . Min-max normalization maps a value v of A to v' in the range $[\text{new_min}A, \text{new_max}A]$ by using the formula given below.

$$v' = \frac{(v - \min A)}{(\max A - \min A)} (\text{new_max } A - \text{new_min } A) + \text{new_min } A$$

In this work, $\text{new_min}A$ was taken as 0 and $\text{new_max}A$ was taken as 1. When min-max normalization is applied, each attribute will lie within the new range of values. However, the underlying distributions of the corresponding attributes within the new range of values will still remain the same. Min-max normalization has the advantage of preserving exactly all the relationships in the data, and it does not introduce any bias.

2.7.5 Feature selection

The main idea of feature selection is to choose a subset of features and eliminate the irrelevant ones from the dataset. Obtaining a smaller set of representative features and retaining the optimal salient characteristics of the data not only decreases the processing time but also leads to more compactness of the models learned, better generalization, and comprehensibility of the mined results. There are two basic feature subset selection techniques - wrapper methods, in which the features are selected using the classifier, and filter methods, in which the selection of features is independent of the classifier used.

-
- **Wrapper methods:** The wrapper method uses a classifier to perform classification using every possible feature subset and finally selects the feature subset that gives the best classification accuracy. This method is deemed to be the best way to improve a classifier's performance since both learning a classifier and selecting features use the same bias (John et al., 1994; Aha et al., 1998). Since a classifier is built as many times as the number of feature subsets generated, computationally intensive classifiers should not be considered. Moreover, it is not advisable to learn a classifier using features selected by another classifier, since different classifiers have their unique biases (Liu et al., 1998b).
 - **Filter methods:** In filter methods, features are first evaluated using feature evaluators. Some popular evaluators are Correlation-based Feature Selection (CFS), Chi-Squared method (CS), Information Gain (IG), Gain Ratio (GR), Symmetrical Uncertainty (SymU), and ReliefF methods. Once the features are assigned scores based on particular criteria such as correlation, chi-square statistic or information measures, they are ranked using methods like best-first, ranker, greedy stepwise etc. In this work, the IG, GR, and SymU methods were chosen because of their ability to pick features that have the minimum entropy and, therefore, the maximum information. The CFS and CS methods implement commonly used statistics - the correlation coefficient and chi-square statistic, respectively, and, hence, were chosen. ReliefF was chosen because it has been shown to work in the presence of noise and inter-feature dependencies to pick the best features based on the nearest neighbors (Kira et al., 1992).

2.7.6 Classifier development and evaluation: Classifiers used

Artificial intelligence is usually defined as the science of making computers do intelligent activities such as learning, reasoning, and problem solving the way humans do. One of the most popular tasks done by these artificial intelligence based computers is classification. Classifiers use features of the input data to learn to classify the data to belong to a particular class. For many years, classifiers have been studied and used in the medical field to improve the prediction accuracy. A review of some of these studies can be found in Paliwal et al. (2009).

Classifier performance depends greatly on the characteristics of the data to be classified. There is no single classifier that works best on all given problems. Hence, in this study, several classifiers were evaluated to understand which classifier performed effectively. Usually for classification tasks, supervised learning is preferred. In this type of learning, during the training phase, both the features (input) and the corresponding class label (output) are presented to help the classifier learn the relationship between the input and output. After training, only the features of the test data are presented. The classifier then automatically predicts the unknown class label with the help of the knowledge gained during the training phase. Unsupervised learning techniques like clustering and self-organizing maps are preferred when the objective is to discover the natural structure of the input data. The most common classifiers (based on supervised learning) used for medical diagnosis are statistical inference techniques and artificial neural networks.

Back Propagation Neural Network (BPNN): The back propagation algorithm consists of two paths: the forward path and the backward path. The forward path deals with creation of a feed forward network, weight initialization, simulation, and training of the network. The network weights and biases are updated in the backward path. During training, the weights and biases of the network are iteratively adjusted to minimize the network performance function (mean square error between network outputs and the target output, which is the class). The weight update aims at maximizing the rate of error reduction, and hence, it is termed as the *gradient descent* algorithm (Wasserman, 1993).

Probabilistic Neural Network (PNN): Probabilistic neural networks are a kind of two-layer radial basis network suitable for classification problems. When an input is presented, the first layer (radial basis layer) computes distances from the input vector to the training input vectors and produces a distance vector whose elements indicate how close the input is to a training input. The second layer (competitive layer) sums these contributions for each class of inputs to produce a vector of probabilities as its net output. Then the *compete* transfer function on the output of the second layer picks the maximum of these probabilities, and assigns a class label 1 for that class and a 0 for the other classes.

Linear Discriminant Analysis (LDA): LDA is used to find the linear combination of features which best separates the two classes.

Support Vector Machines (SVM): Support vector machines are another set of related supervised learning methods used for classification. An SVM will construct a separating hyperplane that maximizes the margin between the input data classes which are viewed in an n -dimensional space (n stands for the number of features used as inputs). To calculate the margin, two parallel hyperplanes are constructed, one on each side of the separating hyperplane. These two hyperplanes are computed directly using the training set.

Fuzzy Inference System (FIS): Fuzzy logic is a form of multi-valued logic to deal with reasoning that is approximate and that is not constrained to belong to precise binary values of either true or false. It can be used to take fuzzy or imprecise observations for inputs and yet arrive at crisp and precise values for outputs. In this work, a Fuzzy Inference System (FIS) was generated with the help of subtractive clustering technique. This clustering technique is used to estimate the number of clusters and the cluster centers in the examined dataset. A FIS is composed of inputs, outputs, and a set of rules that dictate the behavior of the fuzzy system. Each input and output has as many membership functions as the number of clusters that was chosen by the clustering technique. A radius parameter is used to specify a cluster center's range of influence in each of the data dimensions. Once the training is over, a FIS structure that contains a set of fuzzy rules to cover the feature space is generated. This is utilized to perform fuzzy inference calculations of the test data.

2.7.7 Classifier development and evaluation: Data resampling technique

Usually, prior to classification or feature selection, the available dataset would be divided into training and test sets (hold-out technique). The most widely used split-up ratio is 70% train set and 30% test set. The training set would be used to select the features, and design the classifier, and the test set would be used to estimate the prediction accuracy of the classifier. This strategy is based on the assumption that the training and test sets are chosen as representatives of the same unknown distribution of

the data. This is usually true for large datasets, but the strategy can result in less generalization and unreliable accuracy in the case of smaller datasets.

When the total number of samples is relatively small, there is a need to go for better resampling techniques such as resubstitution method, leave-one-out method, k -fold cross validation, and bootstrap techniques. The most popular and time-tested efficient technique for small sample size (at least greater than 50 samples) is the k -fold cross validation technique. This method divides the available samples into k approximately equally sized disjoint subsets. $(k-1)$ subsets are used for training, and the remaining subset is used for testing to get the performance measures. The process is repeated k times and the final performance measure is the average of all the k measures. In classification problems, stratified k -fold cross validation is used. This technique is similar to cross validation, with one important addition: the subsets are built so as to preserve the original class distribution in all subsets, i.e., in this work, the proportion of benign to malignant samples is the same in all subsets. This guarantees that both the training and test sets have approximately equal class distribution.

2.7.8 Software used

For the data pre-processing steps, *SPSS version 15.0*, *Matlab version 7.4.0.287 (R2007a)*, and *Microsoft Excel* were used. For feature selection, *MedCalc version 9.3.7.0* (<http://www.medcalc.be/>) and *WEKA version 3.6.0* (Hall et al., 2009) were used in this work. Details of the filter feature selection techniques employed in WEKA can be found at <http://weka.sourceforge.net/doc/>. For classifier development and evaluation, Matlab was used. To extract association rules, the software *ARMADA data mining tool version 1.3.2* (<http://www.aiai.ed.ac.uk/~jmalone/software.htm>) was interfaced with Matlab, and used.

2.8 Association rule mining based classification - a review

Association Rule Mining (ARM) is a data mining technique that is used to extract correlations between attributes (features) in a database. These correlations are in the form of rules. Consider an association rule of the following form

$$X \Rightarrow Y$$

where

X is the antecedent / Left-Hand Side (LHS) of the rule. X is of the form $\{(F1: \text{value } 1) \text{ and } (F2: \text{value } 2)\}$, where $(F1: \text{value } 1)$ is called an attribute-value pair or an 'item'. Thus, X is a collection of items, or in other words, X is an itemset.

Y is the consequent / Right-Hand Side (RHS) of the rule, which takes up a form similar to that of X.

If only the class feature is considered in the rule consequent (e.g., Y is $(c: 2)$), then the rules are called classification rules; else, the rules are called association rules. In ARM based classifiers, also called as associative classifiers, generally classification rules are extracted and used for classification. These rules are characterized by two important measures called support and confidence. Support is defined as the proportion of transactions in the dataset that contain X and Y. A transaction denotes a row in a database. Confidence of the rule is defined as the ratio of the support to the number of transactions that contain only X.

The three main phases in developing an associative classifier are: (1) rule generation from the training data, (2) rule pruning to remove irrelevant rules, and (3) classification of new test data using the selected rules. In step one, association rules are generated by applying association rule mining algorithms on a training dataset. This dataset, traditionally, consists of n samples, each sample described by m attributes or features and a class attribute c . The attributes may be categorical or continuous. Values of a categorical attribute are mapped to positive integers. Values of a continuous attribute are discretized into intervals, and each interval is coded into a positive integer. The most commonly used discretization algorithm is the one proposed by Fayyad & Irani (1993). Mostly, only classification rules are selected from the initial rule base, i.e. all the rules will have a class label on the RHS. In step two, the noisy and redundant rules are removed using various pruning strategies. The pruning techniques find a smaller subset of rules that best cover the training dataset. In step three, the selected rules are used to classify new data. A classification rule $X \Rightarrow c_l$ is said to correctly classify a sample n_l , if X is present in the row describing n_l (denoted as a match), and the class label of n_l and the class label of the rule are equal to c_l . The performance efficiency of the developed classifier is quantified by using the accuracy

measure. The accuracy is the fraction of the number of correctly classified test samples over the total number of tested samples. In this section, the details of a few popular rule generation algorithms, rule pruning techniques, and associative classifiers are given.

2.8.1 Algorithms for rule generation

In general, there are two steps involved in rule generation: frequent itemset counting and rule generation. A frequent itemset is an itemset with support greater than or equal to a minimum support threshold. The first step counts the frequency of itemsets and finds frequent ones. The second step generates high confidence rules from the selected frequent itemsets. It is to be noted that if there are q items in a database, there will be a total of 2^q possible candidate itemsets. Therefore, the complexity of the first step is exponential, and hence, many studies that have been carried out focused on finding efficient frequent itemset search strategies (FIMI repository: <http://fimi.cs.helsinki.fi/>). Some algorithms are briefly described in this section.

Apriori algorithm

In the popular Apriori algorithm introduced by Agrawal et al. (1993, 1994), frequent itemsets or patterns are extracted in levels. At each level, the frequent itemsets generated in the previous level are used to form new candidate itemsets. The following is a step-wise approach to this method, assuming a minimum support value (min_support) and a minimum confidence value (min_conf).

1. Initializations: Level = 1, $k = 1$.
2. Find a set of candidate itemsets C_p , which is a set of p -itemsets. For example, 1-itemsets denotes one attribute-value pair, like (F1: 3). 2-itemsets denotes (F1: 3, F2: 4).
3. Scan the dataset to determine the support of each p -itemset. Select only those p -itemsets which have support $\geq \text{min_support}$, and form a set of frequent itemsets L_p .
4. Level = Level + 1; $k = k + 1$;
5. Join the items in L_{p-1} , to generate a new candidate itemset C_p .

6. Scan the dataset to determine the support of each p -itemset in C_p . Select only those p -itemsets which have support $\geq \text{min_support}$, and form a frequent itemset L_p .
7. If the L_p is not null, repeat steps 4 to 6, else go to step 8.
8. For each frequent itemset f_i selected in the entire process, generate non-empty subsets of f_i .
9. Select any non-empty subset s of f_i to form rules of the format " $s \Rightarrow (f_i - s)$ ", if it has a confidence $\geq \text{min_conf}$. Here, the minus sign indicates "exclusion of s from f_i ".

In step 8, the Apriori algorithm uses the "downward-closure" property of support; that is, if an itemset is frequent, then all of its subsets are frequent as well. This also means that if an itemset is infrequent at a level, it can be removed because any addition of items to it in subsequent levels will not make it frequent. The Apriori algorithm performs a breadth-first search in the search space. The limitation of the Apriori algorithm is the time taken to find frequent itemsets from all possible candidate itemsets at each level. Thus, when the size of the dataset is huge, the number of itemsets is high and hence, the Apriori algorithm takes up lot of time and memory. Moreover, with long frequent patterns and low min_support thresholds, repeated database scanning and pattern matching become tedious.

Frequent pattern tree

To alleviate the time and space complexity associated with the Apriori algorithm, the step of candidate generation was omitted in the Frequent Pattern tree (FP-tree) approach (Han et al., 2000; Han et al., 2004). In this method, an extended prefix-tree structure that stores information about the frequent patterns is constructed. The nodes in this tree are formed by frequent length one items (items of length one which satisfy an absolute support threshold). As a result, a compact tree, which depicts all the frequent itemsets in each row of the database, is made available, and subsequent mining is done from this tree rather than from the whole dataset. Moreover, only two database scans are required: one for generating all the frequent itemsets, and the second, for constructing the tree. Once the FP-tree is built, a pattern fragment growth method is used for forming rules. For each frequent pattern, all possible combinations of patterns are generated and stored in a conditional FP-tree with the help of node

links. This method then mines all the frequent itemsets by recursively finding all frequent one-itemsets in the conditional FP-tree. This approach of decomposing the mining problem into a set of smaller problems is called the divide-and-conquer approach. This approach was found to be at least an order of magnitude faster than the Apriori algorithm. One limitation that was observed with this approach was that the FP-tree does not always fit into the main memory.

ECLAT algorithm

ECLAT is the first algorithm to use depth-first search strategy for frequent pattern mining. Generally, the dataset used can be represented in two ways: horizontal layout, where each transaction has an identifier (id) followed by the list of features for that transaction; vertical layout, where the dataset consists of a group of items and each item is followed by a tid-list. A tid-list of an item is the list of all transaction ids (tids) that contain that item. ECLAT determines the support by using the intersection of tids. It is based on equivalence class clustering (Zaki et al., 1997). In an updated algorithm called *dEclat*, Zaki et al. (2003) use the difference of tids between the candidate p -itemset and the previous $(p-1)$ frequent itemsets, instead of using the intersection of tids. This algorithm performed better for non-sparse databases.

There are many other algorithms in the literature. Castelo et al. (2001) developed the MAMBO algorithm, which extracted association rules based on conditional independencies. Pasquier et al. (1999) suggested that it would be sufficient to mine only those itemsets that do not have any immediate supersets with the same support (called the closed frequent itemsets). Based on this concept, Pei et al. (2000) and Zaki and Hsiao (2002) proposed efficient algorithms called CLOSET and CHARM, respectively.

2.8.2 Rule pruning techniques

The initial rule set, obtained using the algorithms described in the previous section, is comprehensive. This is, in theory, an advantage for an associative classifier, as it will not miss potentially important and informative rules. However, in practice, the number of rules is prohibitively large and hence, implementing a classifier with such a huge rule base would affect its performance in terms of time and space. Due to this,

rule pruning techniques have been developed to determine a manageable rule base with as low information loss as possible. It is necessary to present a few relevant definitions here, before describing the pruning techniques. Let I denote the initial rule set. Let D be the set of rules pruned from I , and P be the resultant pruned rule set ($P = I - D$).

- **Rule ranking:** A few pruning techniques prune rules according to the rank of the rules. The most common rule ranking or sorting technique is based on confidence, support, and length parameters. According to this technique, given two rules A and B , A precedes B , if the confidence of A is greater than that of B . When both have similar confidence values, A precedes B , if the support of A is higher than that of B . In worst cases, where both confidence and support values of both A and B are equal, A precedes B , if it has lower antecedent cardinality. Thabtah et al. (2004) proposed an additional parameter for rule ranking - the class distribution frequency. If two rules have similar confidence, support, and cardinality, A precedes B , if it is associated with a more representative class than that of B .
- **Idempotency:** A pruning technique satisfies this property if it provides the same final pruned dataset P , irrespective of any number of its application to I .
- **Transparency:** If a classifier built on P gives the same accuracy as a classifier built on I , then the pruning technique is said to be transparent. Transparent techniques can be used to obtain a reduced rule base without compromising on the accuracy of classification.
- **General and special rules:** Let A and B be two rules. $A: X \Rightarrow c$ and $B: Y \Rightarrow c$, where X and Y are itemsets and c is the class. A is said to be a general rule of B (and B is called the specialization of A), if (1) A has lower antecedent cardinality than B , and (2) both have the same class label.

Few popular rule pruning techniques are described below. Rule pruning may be carried out at early stages, for instance, removing rules that do not pass a minimum support or confidence. Or, it may be carried out while generating the rules (chi-square method), or at later stages (database coverage).

Redundant rule pruning

It is evident that a general rule A matches a superset of the data covered by any of its specialization rules, since the antecedent of A has fewer items than that of its specialization rules. Also, general rules and their specific rules will classify the samples in a similar manner, as they have the same class labels, thereby maintaining the accuracy. Moreover, A will have a higher rank compared to its specialization counterparts because of its lower antecedent cardinality. For these reasons, the redundant rule pruning technique aims at removing specialization rules, if their general rules are present. This technique satisfies both idempotency and transparency properties. This pruning technique is based on the associative rule cover approach presented in Toivonen et al. (1995). Associative classification algorithms presented in Li et al. (2001), Antonie et al. (2003), and Antonie et al. (2004) use this technique.

Confidence based pruning

This is a simple pruning approach wherein rules with confidence above a given threshold are selected. This process is already a part of the Apriori algorithm and has been used in many algorithms developed by Baralis et al. (2008), Li et al. (2001), Liu et al. (1998c), and Liu et al. (2000). This pruning technique satisfies the idempotency property, but does not satisfy the transparency property as accuracy might not be the same using the pruned rule set.

Database coverage

The database coverage technique was introduced in Liu et al. (1998c). In this technique, rules are first generated using the training dataset and then ranked. Starting from the rule A with the highest rank, a pass is made over the training dataset to select the samples that match with the LHS of A . The rule A is placed in P , and the samples matching A are removed from the training set. The procedure is repeated until all training samples in the database are covered, or all ranked rules are considered. If a rule does not correctly cover at least a single sample, it is discarded. This technique has also been employed in Liu et al. (2000), Li et al. (2001), and Thabtah et al. (2004). The idempotency property is satisfied by this technique, whereas it is not transparent. The main limitation of this technique is that the pruned rule set P is strongly customized to the training set, and hence, the higher the discrepancy between the characteristics of the test set and that of the training set, the lower will be the accuracy

of the classifier using P . As an improvement to this technique, Li et al. (2001) proposed that a sample be discarded from the training set only after it has been covered by a user-specified minimum number of rules, instead of a single rule.

Chi-Square (χ^2) test based pruning

The chi-square test is a widely used statistical technique which evaluates the correlation between two variables and determines whether they are independent or correlated (Mills, 1965). In the case of association rule mining, this test is used to find the correlation between X and c in the case of a rule $X \Rightarrow c$. If they are positively correlated (rule passes the chi-square test), the rule is placed in P . The significance level that is usually considered to check the correlation is 3.84% (Baralis et al., 2008; Li et al., 2001). This technique satisfies idempotency but not transparency. The CMAR algorithm adopts this technique in the rule extraction phase itself. Only rules satisfying the chi-square test are extracted.

Pessimistic Error Rate (PER) based pruning

This technique was initially proposed for decision trees (Quinlan, 1993). In the context of associative classification, PER is defined as follows

$$PER(A) = \frac{E}{N} + \frac{1}{2N}$$

where E is the number of training samples correctly classified by the rule A , and N is the number of samples matched by the rule A . The term $\frac{1}{2N}$ is a continuity correction factor for binomial distribution (Quinlan, 1993).

If the PER value of A is higher than the PER value of at least one of the general rules of A , A is discarded. This method was proposed by Liu et al. (1998c). This technique satisfies idempotency but not transparency.

Lazy pruning

In this technique, only the rules that lead to incorrect classifications are discarded from the rule set. The first rule in the ranked rule list that matches a training sample is

assigned to the sample, and its classification correctness is checked. If correct, the training object is removed and the rule is kept. Once all the samples are considered, the rules that resulted in wrong classification are removed, and the samples covered by these rules are subjected to another cycle. This process is continued until all the training samples are correctly classified. The resulting rule base has one set of rules, which classified at least one sample, and another set of unused rules. Since this pruning technique is less customized to the training set, it might handle new data better when compared to the database coverage method. However, the number of rules generated is much higher than the database coverage technique, which makes it difficult for interpretation. This technique satisfies idempotency but not transparency. It has been implemented in Baralis et al. (2002) and Baralis et al. (2004).

Conflicting rules

In some cases, it is possible to extract rules that have the same LHS, but different RHS or class labels. Antonie et al. (2003) suggested removing these conflicting rules from the rule set. However, Thabtah et al. (2004) indicated that such rules might represent useful knowledge, and hence, have proposed that these rules be combined into one multi-label rule (MMAC algorithm).

2.8.3 Ways to classify test samples

In associative classifiers, the class of a test sample may be predicted by using a single highest precedence rule applicable to that sample, or by using multiple rules.

Maximum likelihood based prediction

In this method, only a single high precedence rule is used for classifying a test sample. If there are no rules in P that match the given test sample, then the dominant class label in the training data set is assigned. This method is simple and is widely accepted, because, in general, the highest ranking rules (with high confidence) may be considered as the rules that indicate the likelihood that a test sample belongs to the right class. However, there is a definite possibility, especially in huge datasets, that there could be more than one rule with similar confidence that match the test sample. Also, the highest confidence rule may be ineffective in databases with unbalanced distribution of class labels (Li et al., 2001; Liu et al., 2000). Thus, it has been

suggested that a subset of rules be used to predict the class (Li et al., 2001; Zaiane et al., 2002).

Multiple rule based prediction methods

In CMAR algorithm by Li et al. (2001), a subset of high confidence rules applicable to a test sample is selected. If all the rules in the subset predict the same class, that class is assigned to the test sample. Else, the algorithm splits the rules into groups based on the class label, and the strength of each group is determined using weighted chi-square analysis. The class belonging to the group with the highest strength is assigned to the sample.

In another algorithm WAR-BC (Weighted Association Rule Based Classification) proposed by Dua et al. (2009), a rule weighting scheme is used. They determined a horizontal weight for each rule, which is based on the frequency of a rule across different classes. They also determined a vertical weight, wherein, the class level rules are first arranged in the decreasing order of confidence-support, and are then assigned a weight equivalent to the number of rules present in the class. If there are 10 rules present in a class, the rule with rank 1 is assigned a weight of 10; a rule with rank 2 is assigned a weight of 9 and so on. An overall score is calculated using these two weights. During classification, the scores of all the rules matching the sample are combined class-wise. The class with maximum score is assigned to the sample.

2.8.4 Various associative classifiers

The first classifier using association rules was CBA (Classification Based on Associations), developed by Liu et al. (1998c). Many other associative classifiers have been developed since then. Table 2.9 presents a summary of some of these classifiers highlighting the algorithm used for rule generation, ranking, and pruning.

Table 2.9 Summary of a few associative classifiers

Name of the associative classifier	Data layout	Rule discovery algorithm used	Rule ranking technique used	Rule pruning techniques used	Reference
CBA	Horizontal	Apriori	Support, confidence, rule cardinality	PER, database coverage	Liu et al. (1998c)
CBA-2 using multiple class supports	Horizontal	Apriori	Support, confidence, rule cardinality	PER, database coverage	Liu et al. (2000)
CMAR	Horizontal	FP-growth	Support, confidence, rule cardinality	Chi-Square, database coverage, redundant rule	Li et al. (2001)
L ³	Horizontal	FP-growth	Support, confidence, rule cardinality, items lexicographical	Lazy pruning	Baralis et al. (2002)
ARC_AC	Horizontal	Apriori	Support, confidence	Redundant rule	Zaiane et al. (2002)
ARC_BC	Horizontal	Apriori	Support, confidence, rule cardinality	Redundant rule	Antonie at al. (2003)
L ³ G	Horizontal	FP-growth	Support, confidence, rule cardinality, items lexicographical	Lazy pruning	Baralis et al. (2004)
MMAC	Vertical	Tid-list intersections and recursive learning	Support, confidence, rule cardinality, class distribution frequency	Database coverage	Thabtah et al. (2004)

CHAPTER 3: BDS CLINICAL STUDY METHODOLOGY

The clinical study was conducted at Tan Tock Seng hospital, after receiving the approval from the Institutional Review Board and ethics committee. A qualified trainer from Philippines was invited, and a three day training workshop was conducted wherein the technique of using BDS was studied. In this chapter, the inclusion and exclusion criteria used for selecting the volunteers for the BDS test, and the clinical study methodology are described.

3.1 Inclusion criteria

Women who were at least 21 years old, who were capable of giving informed consent, and women who had a suspicious breast lesion located by palpation and/or imaging technology, were considered as eligible for the test. Just for evaluation purposes, a few asymptomatic women were also examined. Since BDS is an adjunct modality that depends on physical examination or ultrasound or mammography results to determine the prior LOS, the patient should also be scheduled to take up one or more of these tests. Even though the BDS specifications limit the upper age to 55 years, the upper limit was set to 80 years for evaluation purposes.

3.2 Exclusion criteria

Efficacy of BDS in patients with the following conditions has not been fully evaluated in previous Biofield research. Hence, women with these conditions were not selected for the test.

- Self-history of breast cancer.
- Underwent chemotherapy or radiation therapy for any cancer within past three years, as these may alter the carcinogenic process and, therefore, affect the BDS test result.
- Underwent fine needle aspiration of a suspicious lesion performed within 24 hours prior to the BDS test.
- Breast augmentations.

Patients with the following conditions were excluded from the study as these conditions are believed to confound the BDS test results.

-
- Underwent active tamoxifen or raloxifene therapy within the past three months as they may alter the carcinogenic process and, therefore, affect the BDS test result.
 - Clinically diagnosed active mastitis as this may produce areas of focal proliferation and confound the test result.
 - Visible or known dermatological condition or other physical abnormality precluding the application of sensors.
 - Allergy to adhesives or gels used to affix ECG or electroencephalogram sensors.
 - Only one breast.
 - Active electrically powered implanted devices (e.g., pacemakers, automatic implantable cardioverter-defibrillator, transcutaneous electrical nerve stimulators, implanted pumps, etc.).
 - Breast tattoos and nipple jewelry.
 - Pregnant women.

3.3 Clinical study methodology

The clinical study methodology employed is listed in steps below:

1. Eligible patients who satisfied the inclusion and exclusion criteria were identified by the doctors and radiographers.
2. The *participant information sheet* (see Appendix B) was given to the selected patient. In the presence of a witness, who could either be the radiologist or the patient's relative, the BDS test procedure was briefed to the patient. If necessary, translators were used to explain the procedure to the woman in the language that she understood. If the patient agreed to take up the BDS test, she was asked to sign the *participant consent form* (see Appendix B).
3. The demographic details such as age, race, menstrual history, pregnancy history, family history of cancer, and previous breast surgery/biopsy details were recorded in a *participant details form* (see Appendix B).
4. The patient was asked to lie in the supine position with the head slightly elevated for comfort. The suspicious breast lump was located by palpation or with the help of her mammograms or breast ultrasound scan images if the lump was non-palpable. The BDS single-use sensor pads were attached to the sensor leads and were then placed on the breasts in the manner described in section 2.4.6.5.

5. Details such as the patient name, an identification number, date of birth, the number of days since the first day of last menstrual cycle, the average number of days between menstrual cycles, and the location of the tested lesion (right or left breast), were input into the BDS device. Subsequently, the system prompts the user to re-check the placement of sensors. Once confirmation was given, the device starts the five minute stabilization period timer. This period is required before actual electropotential measurement happens. Fifteen seconds before the data acquisition begins, the device will beep, indicating that the test is about to begin and that the patient be quiet and still.
6. After a few minutes, data collection will be completed and the BDS index is calculated by the system. If the test was successful, the device presents the BDS index and prints the same on a *test result form* (see Appendix B). If the test was unsuccessful, the device indicates the leads whose data were found to be invalid, and asks the operator to check them and repeat the test.
7. Once the BDS test was completed, the leads were removed from the sensor pads. The sensor pads were then removed from the patient and discarded. The gel on the skin was wiped off using a tissue.
8. The patient feedback on the ease and the comfort of the test when compared to mammogram and ultrasound was then obtained.
9. The prior LOS of the patient was determined by the doctor using CBE, ultrasound or mammography results. As highlighted earlier, this LOS score can be 1-normal, 2-benign tumor or cyst/ microcalcifications, 3- malignancy cannot be excluded, 4- probable malignancy and 5-malignant.
10. The obtained BDS index and the prior LOS value were used in the interpretation *Map 2* (Table 2.4) to determine the Post-BDS LOS of the patient. This Post-BDS LOS represents the final diagnosis made by the BDS test.
11. For statistical analysis and classification, the actual diagnosis of the patient (whether she has cancer or not) should be known. This was obtained using her biopsy results. If the biopsy was not done, it means the doctor was confident enough to rule it out. So the doctor's diagnostic prediction was taken as the actual disease status for that patient.

According to the literature, the BDS index of the cancerous lesions is significantly greater than that of the benign lesions in premenopausal women after the third day of

the menstrual cycle and also when the test is performed after 10 A.M. (Cuzick et al., 2002; Sacchini et al., 2002). Hence, these key parameters were considered during patient recruitment. As per the device protocol, the patient recruitment procedure cannot be blinded, i.e., the knowledge of whether the patient is “probably benign” / “probably malignant” should be known. Even though this information is not needed for the calculation of the BDS index, it is necessary for the deduction of prior LOS of the patient and therefore, the Post-BDS LOS. This is acceptable, since BDS is currently approved to be an adjunct modality and it is designed to revise the existing LOS of the patient.

The BDS test forms, which were obtained from each patient, were stored securely at the hospital to prevent any breach of patient confidentiality. Since most patients were not scheduled for immediate biopsy, follow up was done at periodic intervals to record the biopsy results.

CHAPTER 4: STATISTICAL ANALYSIS OF THE BDS DATASETS

Two BDS datasets were available for further analysis. One was the TTSH dataset and the other was the US dataset obtained from the manufacturer. The US dataset was acquired during the clinical trials conducted by BDS in the United States. The effect of demographic parameters such as age, menstrual history, pregnancy history, and family history of cancer, and the effect of lesion based parameters such as size, location, and palpability status on the cancer incidence are analyzed in this chapter. The sensitivity, specificity, accuracy, and the AUC measures of BDS are also presented.

4.1 Analysis of the TTSH dataset

4.1.1 TTSH dataset description and pre-processing

195 women scheduled for mammography and/or ultrasound tests participated in the BDS clinical study. The feedback given by each participant on the comfort of the procedure was positive. Adverse events that happened during the test included only one sensor misplacement. One patient experienced mild transient anxiety in general. However, no BDS evaluations were interrupted or discontinued due to an adverse event. Among the 195 cases, 53 were malignant, 96 were benign, 33 were normal, one patient withdrew from the study, eight cases were microcalcifications, three cases had to be discarded due to variability between the tested and biopsied lesion location, and finally, for one suspicious case, the biopsy results were unavailable. Considering only the malignant, benign and normal cases, there were a total of 182 cases for analysis. Of the 182 cases, 53 cases were malignant (age = 55.4 ± 9.2 years), 96 were benign (age = 44.8 ± 9.7), and 33 were normal (age = 51.1 ± 11.2). Among the 96 benign cases, 48 cases were biopsy proven. The remaining 48 cases were visibly or radiologically declared benign. All the malignant cases were declared positive using biopsy. Overall, 101 out of 182 cases were biopsy proven. The percentage breakdown of the total number of cases is summarized in Table 4.1.

The following features were available in the TTSH dataset: age, race, menarche age, menopause age, pregnancy history (age of first full term pregnancy or nulliparous),

Statistical analysis of the BDS datasets

family history, lesion palpability status, lesion location, lesion size, hormonal usage, breast biopsy history, lifestyle factors like smoking and drinking, mammography LOS, ultrasound LOS, prior LOS, BDS index, Post-BDS LOS, and class (benign/malignant). The class represents the actual diagnosis for the patient. Before classification, the data has to be pre-processed to counter missing data and to remove outliers. In the TTSH dataset, the size parameter was missing in three cases, and hence, the class average of size was used to counter that. No outliers were present. Since the tested BDS device was meant for clinical purposes, it did not output the 16 sensed electropotential values.

Table 4.1 TTSH dataset: Percentage breakdown of the total number of cases

	Total	Percentage
Number of discarded cases	4	2.05
Number of microcalcification cases	8	4.10
Tests with malignant diagnoses	53	27.18
Tests with benign diagnoses	96	49.23
Tests with normal diagnoses	33	16.92
Tests with indeterminate results	1	0.5
Total tests with case reports	195	100

4.1.2 TTSH dataset: Study of effect of risk factors on cancer incidence

The patient demographic data were analyzed. This section presents the analysis of the effect of risk factors on the cancer prevalence among patients.

Menstrual history: Out of the 81 patients who had attained menarche before the age of 13, 21 had cancer. That is, there is 25.9% of cancer incidence among women who had menarche before the age of 13. In another perspective, 21 out of 53 malignant cases (39.6%) occurred in women with this risk factor. Among the 63 women who menopaused after 50 years of age, 31 malignant cases were observed, indicating 49.2% of cancer incidence. In another perspective, 31 out of 53 malignant cases (58.5%) occurred in women with this risk factor. Among the 83 post-menopausal women, 33 had cancer, i.e. 39.8% of post-menopausal women had cancer. In another perspective, 33 out of 53 cancers (62.3%) occurred in post-menopausal women.

Pregnancy history: Only three out of 53 malignant cases (5.7%) occurred in women who had their first full term pregnancy after 30 years, and 24.5% (13/53) of cancers appeared in women who were nulliparous.

Family history: Only seven out of the 24 women with family history had cancer.

Age: Age is a vital risk factor for breast cancer. The diagnostic results, split according to the age group, are given in Table 4.2. Out of the 53 women diagnosed with cancer, 40 (75.5%) were above 50 years of age. Put in another way, 40 out of the 87 women aged 50 or above (46%) had cancer. This proves that age is an important risk factor for the incidence of cancer.

Table 4.2 TTSH dataset: Age distribution and respective diagnoses

Age range	No. of patients
Between 21 and 30 (both ages inclusive)	10
Normal	1
Benign (4 biopsied, 5 unbiopsied)	9
Malignant	0
Between 31 and 40 (both ages inclusive)	22
Normal	6
Benign (7 biopsied, 8 unbiopsied)	15
Malignant	1
Between 41 and 49 (both ages inclusive)	63
Normal	5
Benign (21 biopsied, 25 unbiopsied)	46
Malignant	12
Greater than 50	87
Normal	21
Benign (16 biopsied, 10 unbiopsied)	26
Malignant	40

The effects of few other risk factors such as previous breast biopsy, usage of OCP and HRT, smoking or drinking habits, and race, were not prominent in this study as they did not have a significant impact on the malignancy incidence.

4.1.3 TTSH dataset: Study of effect of lesion parameters on cancer incidence

This section presents the analysis of the effect of lesion parameters on the cancer prevalence among patients.

Lesion location: Each breast can be divided into nine regions - upper outer quadrant, upper inner, lower outer, lower inner, 3 o'clock, 6 o'clock, 9 o'clock, 12 o'clock positions, and behind the nipple (subareolar). Table 4.3 presents the lesion distribution at these positions in the studied population. The table also lists the location of

Statistical analysis of the BDS datasets

malignant lesions. A total of 149 lesions (53 malignant and 96 benign) were examined. Most patients had lesions at positions other than the examined lesion position. However, only the most suspicious lesion in each patient was examined. The remaining 33 patients had normal breasts without any lesions.

Table 4.3 TTSH dataset: Location of lesions

	Location	No. of lesions	Percentage (%)	No. of malignant lesions	Malignant lesion percentage (%)
Upper outer quadrant	Right	24	36.2	9	34
	Left	30		9	
Upper inner quadrant	Right	15	21.5	7	26.4
	Left	17		7	
Lower outer quadrant	Right	9	8.1	5	11.3
	Left	3		1	
Lower inner quadrant	Right	8	8.1	3	5.7
	Left	4		0	
Other locations	Right 3 o'clock position	3	26.2	1	22.6
	Right 6 o'clock position	2		1	
	Right 9 o'clock position	13		3	
	Right 12 o'clock position	3		1	
	Right subareolar	3		0	
	Left 3 o'clock position	6		2	
	Left 6 o'clock position	3		0	
	Left 9 o'clock position	0		0	
	Left 12 o'clock position	4		3	
	Left subareolar	2		1	
Total		149	100	53	100

Out of the 149 examined lesions, 36.2% of the lesions appeared in the upper outer quadrant of both breasts and 21.5% in the upper inner quadrant. The high presence of lesions in the upper quadrant can be attributed to the increased amount of glandular tissue present in this area. An interesting observation was that around 81% of the malignant lesions belonged to upper outer, upper inner, subareolar, and 3, 9, and 12 o'clock positions. This agrees to the generally observed statistics that most of the malignant lesions appear in upper quadrant or behind the nipple (around 55% of the breast cancers are found in the upper quadrant of the breast and 34% in the area behind the nipple (Imaginis: <http://www.imaginis.com/breasthealth/performbse.asp>).

Lesion palpability nature: Among the 149 lesions that were examined, 105 were palpable and 44 were not palpable. Out of the 53 malignant lesions, 47 (88.7%) were palpable.

Lesion size: The lesion size (the longest diameter of the lesion) varied between 3 and 89 millimetres. All the malignant lesions measured at least 6 millimetres along the largest diameter.

4.1.4 TTSH dataset: Results of diagnostic tests

Table 4.4 presents the different possible combination of diagnostic tests and the number of patients who took them.

Table 4.4 TTSH dataset: Patient distribution in terms of diagnostic tests undertaken

Combination of diagnostic tests	No. of patients
Mammography, Ultrasound, BDS, Biopsy	89
Mammography, BDS, Biopsy	1
Ultrasound, BDS, Biopsy	11
Mammography, Ultrasound, BDS	48
Mammography, BDS	14
Ultrasound, BDS	19

There has been no mention in the literature about ultrasound and/or mammogram tests affecting the impedance of the lesions. Hence, the BDS test can be carried out anytime before or after the patient undergoes ultrasound and/or mammogram. However, BDS should be done before the biopsy of the suspicious lesion, as per the exclusion criteria listed earlier in section 3.2. All the women were offered CBE. However, the LOS values for CBE were not available in the TTSH dataset.

For the purpose of getting sensitivity, specificity, accuracy, and AUC values, the 33 normal cases were considered as benign cases. Hence, overall there were 129 benign cases and 53 malignant cases. In the case of ultrasound and mammography, the test result was deciphered from the respective BI-RADS score (BI-RADS score $\leq 3 \Rightarrow$ benign; BI-RADS score $\geq 4 \Rightarrow$ malignant). In the case of BDS, the Post-BDS LOS was used (Post-BDS LOS $\leq 3 \Rightarrow$ benign; Post-BDS LOS $\geq 4 \Rightarrow$ malignant).

Ultrasound results:

Out of the 182 cases, 15 women (14 benign cases and one malignant case) had not taken up the ultrasound test. In order to make a fair comparison of the results between the three tests (ultrasound, mammography, and BDS), the number of ultrasound cases should be adjusted to 182. Assigning the corresponding mammogram LOS as the ultrasound LOS for the patients who did not take up ultrasound is not a good assumption as these two tests are not equal in terms of results. Therefore, the Ultrasound LOS (ULOS) values were assigned to these 15 women based on the following conditions:

- If class was malignant and MG LOS was 4 or 5, ULOS = MG LOS
- If class was benign and MG LOS was 1,2, or 3, ULOS = MG LOS
- If class was malignant and MG LOS was 1, 2, or 3, ULOS = 4.
- If class was benign and MG LOS was 4 or 5, ULOS= 3

Considering all 182 cases, the sensitivity, specificity, and accuracy were 96.23%, 94.57%, and 95.05%, respectively. The contingency table is presented in Table 4.5.

Table 4.5 TTSH dataset: Ultrasound results (N=182)

Ultrasound result	Disease status	
	Present	Absent
Positive	True positive - 51	False positive - 7
Negative	False negative - 2	True negative - 122
Results		
Sensitivity	96.23%	
Specificity	94.57%	
Accuracy	95.05%	

Mammography results:

Out of 32 women under the age of 40, 20 women were not offered mammography because of radiation concerns. Overall, out of the 182 cases, 30 women had not taken up mammography test. For these cases, MG LOS value was assigned based on the following conditions:

- If class was malignant and ULOS was 4 or 5, MG LOS = ULOS
- If class was benign and ULOS was 1,2, or 3, MG LOS = ULOS
- If class was malignant and ULOS was 1,2, or 3, MG LOS = 4
- If class was benign and ULOS was 4 or 5, MG LOS = 3

Statistical analysis of the BDS datasets

The MG test resulted in a sensitivity of 90.57%, specificity of 98.45%, and an accuracy of 96.15%. The contingency table is presented in Table 4.6.

Table 4.6 TTSH dataset: Mammography results (N=182)

	Disease status	
Mammography result	Present	Absent
Positive	True positive - 48	False positive - 2
Negative	False negative - 5	True negative - 127
Results		
Sensitivity	90.57%	
Specificity	98.45%	
Accuracy	96.15%	

BDS results:

The BDS test recorded a sensitivity of 96.23%, specificity of 93.80%, and an accuracy of 94.51%. An interesting observation was that 24 out of 33 normal cases registered a Post-BDS LOS of 2 due to low BDS index. This low BDS index indicates that the normal cases have low surface electropotential differentials. The other nine cases had a Post-BDS LOS of 3. The contingency table is presented in Table 4.7.

Table 4.7 TTSH dataset: BDS results (N=182)

	Disease status	
BDS result	Present	Absent
Positive	True positive - 51	False positive - 8
Negative	False negative - 2	True negative - 121
Results		
Sensitivity	96.23%	
Specificity	93.80%	
Accuracy	94.51%	

If only the 149 cases were considered (omitting the normal cases), the sensitivity, specificity, and accuracy values were 96.23%, 91.67%, and 93.29%, respectively. In one of the false positive cases of the BDS test, the lump was linked to diabetic mastopathy, and hence, was assigned a prior LOS of 3. The BDS index of +0.5 was obtained, which, according to Table 2.4, resulted in a Post-BDS LOS of 4 (probably malignant). However, doctors considered the lump to have a high probability of being benign and did not recommend a biopsy. It is evident that BDS has indicated a false positive result. In a study conducted in the Istituto Nazionale Tumori in Milan, chronic mastitis associated with fibrosclerosis resulted in a false positive result

(Sacchini, 1996). Since in this case, the lump is linked to diabetic mastopathy, which is an uncommon chronic mastitis, it can be concluded as a probable reason for the false positive result.

ROC results:

The ROC curves of BDS, mammography, and ultrasound are illustrated in Figure 4.1. The thresholds that gave the best sensitivity and specificity values and the AUC of these tests are presented in Table 4.8. It is evident from Table 4.8 and Figure 4.1 that all the three diagnostic tests have clinically acceptable significant AUC values. The ROC curves for ultrasound and BDS almost merge resulting in almost similar AUC. The *decision threshold* (LOS value) that provides the optimum value for sensitivity and specificity for each diagnostic test is indicated by an asterisk. Thus, from the ROC results, it is evident that the efficiency of BDS is highly comparable to that of mammography and ultrasound.

Table 4.8 TTSH dataset: Summary of ROC results of ultrasound, mammography, and BDS tests

ROC RESULTS (Sn: Sensitivity; Sp: Specificity; Th: Threshold)								
Ultrasound N =182			Mammography N =182			BDS N=182		
Th	Sn (%)	Sp (%)	Th	Sn (%)	Sp (%)	Th	Sn (%)	Sp (%)
>=1	100	0	>=1	100	0	>=2	100	0
>1	100	18.60	>1	96.23	34.88	>2	100	67.44
>2	100	71.32	>2	92.45	77.52	>3 *	96.23	93.80
>3 *	96.23	94.57	>3 *	90.57	98.45	>4	28.30	100
>4	26.42	100	>4	26.42	100	>5	0	100
>5	0	100	>5	0	100			
Area under the curve		0.975	Area under the curve		0.945	Area under the curve		0.972

Among the three diagnostic tests, mammography recorded the highest accuracy and specificity but lowest sensitivity. This lowered sensitivity may be attributed to the fact that four out of the five false negative cases occurred in young women (between 40 and 52 years of age) with dense breasts, and it is a known fact that mammography is less sensitive for these cases. It may also be because of the lesser number of malignant samples available in this study.

Figure 4.1 TTSH dataset: Comparison of ROC curves of ultrasound, mammography, and BDS tests

4.1.5 TTSH dataset: Biopsy reduction

To know if BDS is helpful in making biopsy reductions, the indeterminate cases with prior BDS LOS of 2 or 3 which have been sent for biopsy have to be considered. One of the key uses of any adjunct modality is to reassure the predictions for these indeterminate lesions. In this dataset, there were 44 such cases, and 42 of these cases had benign biopsies. In 41 of these cases, the BDS test resulted in Post-BDS LOS of either 2 or 3. Even in such a small sample size, BDS test has demonstrated that biopsy could have been avoided for these 41 cases. i.e., BDS has a 93.18% (41/44) accuracy level in predicting unwanted biopsies.

4.2 Analysis of the US dataset

4.2.1 US dataset description and pre-processing

The US dataset was from one of the BDS trials conducted in the United States back in 1995. The dataset was obtained from the manufacturer. It had a total of 291 cases, out of which 198 were benign and 93 were malignant. The major difference between this dataset and the TTSH dataset is the presence of 16 sensor electropotential values for each patient in this dataset. This dataset was purposely chosen to study the effect of

these additional sensor values on the disease prediction efficiency using classifiers (next chapter). Thus, for each patient, the following features were available in the dataset: menopausal status (Pre, Post, and Peri), pregnancy history, family history, lesion palpability, location and size, age, 16 raw measured electropotential values, prior LOS, BDS score, Post-BDS LOS, ultrasound LOS, mammography LOS, CBE LOS, and biopsy result (class). The results of the pre-processing steps employed on this dataset are presented in this section.

Missing data handling: The one case with missing menopause status was assigned the “post” menopausal status, considering her age (45 years). The status “peri” indicates that the women are almost attaining menopause. Hence, the 15 peri-menopausal cases were assigned “post” menopausal status because it is easier to handle the variable with binary discretization (either pre or post). Moreover, the age range of these peri-menopausal cases was between 44 and 57 years (quite a possibility for menopause). Over 30% of the cases did not have the lesion size recorded. Therefore, the parameter was removed from the analysis. The other features did not have any missing values.

Outlier removal: Most of the electropotential features did not have a normal distribution, according to normality plots and Shapiro-Wilk test. One probable reason for non-normality is the presence of outliers. Also, according to the BDS Physician manual, the difference between the electropotentials (SC-SV) of cancers and the electropotentials (SC-SV) of non-cancerous cases should be statistically significant. With outliers, the data did not exhibit this difference (as seen in Table 4.9). Hence, the following techniques were employed to remove the outliers from the dataset.

- **Electropotential value based outliers (for malignant cases):** It has been reported that the *value recorded by the SC sensor for malignant cases must be statistically greater than or equal to the mean of SI, SO, SU, and SL values* (If the value is lesser, it implies that the SC sensor has not been placed near the site of maximum depolarization leading to lesser separation between malignant and benign results) (Cuzick et al. 2002). Taking that as the null hypothesis, Matlab was used for an unpaired comparison of the two means (mean of SC and mean of SI, SO, SU, and SL) using t-test (unequal variance, left tailed). Originally, there is only one value recorded for each of the electropotentials for each case. To apply t-test, the number

of samples was increased by adding and subtracting '0.1' five times from the original value of the electropotential, resulting in ten additional samples for each electropotential. This calculation ensured that the mean and variance of all the additional samples did not deviate much from the original. All the cases that had a p-value less than 0.05 were considered outliers, i.e. the null hypothesis was rejected indicating that the SC value was not greater than or equal to mean of SI, SO, SU, and SL values. There were a total of 64 malignant cases after outlier removal.

- ***Univariate and multivariate outlier removal (both malignant and benign cases):***
The univariate outliers in each of the 16 electropotential features were detected using z-score method. The threshold was kept as ± 3 for benign dataset (since sample size was greater than 80) and as ± 2.5 for malignant dataset (sample size ≤ 80). The multivariate outliers were detected using Mahalanobis distance. Cases having a probability of the distance ≤ 0.001 were considered outliers. At the significance of 0.01, many cases were marked as outliers. For a dataset of this size, removing all these outliers may result in valuable data being lost. Moreover, on scrutinizing these outliers, it was found that they did not have extreme values. Therefore, for better information preservation, a conservative p-value of 0.001 was used.

After these data-preprocessing steps, there were a total of 241 cases. Of the 241 cases, 58 cases were malignant (age = 57.0 ± 13.57 years) and 183 were benign (age = 48.3 ± 11.37). All the cases had biopsy results. These 241 cases are used in the subsequent analyses. Once the outliers were removed, all the features had a normal distribution. Also, the outlier removal step improved the difference between SC and SV significantly for benign and malignant samples, as depicted in Table 4.10. The p-values shown in Tables 4.9 and 4.10 were obtained using Excel (two tailed t-test; unequal variance).

Statistical analysis of the BDS datasets

Table 4.9 Comparison of electropotential differentials before outlier detection

Histological Diagnosis/ Number of Cases	SC <i>Mean of Center sensor @ quadrant with suspicious lesion ± Std error</i>	SV <i>Mean of Vertical sensor in uninvolved quadrant ± Std error</i>	Mean of Differential SC-SV
(1) Malignant/ N=93	9.52 ± 1.36 mV	7.80 ± 1.15 mV	1.71 mV SC vs. SV p=0.3385 (NS)
(2) Benign/ N=198	9.11 ± 0.82 mV p= 0.8024 (NS)(1 vs. 2)	6.94 ± 0.74 mV p= 0.5370 (NS)(1 vs. 2)	2.17 mV SC vs. SV p=0.05 (S) p=0.7057 (NS)(1 vs. 2)

(VS – Very Significant at even 1% level; S- Significant at 5% only; NS – Not Significant)

Table 4.10 Comparison of electropotential differentials after outlier detection

Histological Diagnosis/ Number of Cases	SC <i>Mean of Center sensor @ quadrant with suspicious lesion ± Std error</i>	SV <i>Mean of Vertical sensor in uninvolved quadrant ± Std error</i>	Mean of Differential SC-SV
(1) Malignant/ N=58	13.57 ± 1.23 mV	8.37 ± 1.28 mV	5.20 mV (SC vs. SV) p=0.0042 (VS)
(2) Benign/ N=183	8.99 ± 0.80 mV p= 0.0023 (VS)(1 vs. 2)	6.8 ± 0.71 mV p= 0.2889 (NS)(1 vs. 2)	2.18 mV (SC vs. SV) p=0.04 (S) p=0.018 (VS) (1 vs. 2)

(VS – Very Significant at even 1% level; S- Significant at 5% only; NS – Not Significant)

4.2.2 US dataset: Study of effect of risk factors on cancer incidence

The patient demographic data were analyzed. This section presents the analysis of the effect of risk factors on the cancer prevalence among patients.

Menstrual history: The menarche details were not available in this dataset. Also, menopause age was not specified. Among the 131 post-menopausal women, 44 had cancer, i.e. 33.6% of post-menopausal women had cancer. In another perspective, 44 out of 58 cancers (75.9%) occurred in post-menopausal women.

Pregnancy history: Only six out of 44 women who never were pregnant had cancer.

Family history: 18 out of the 78 women with family history had cancer.

Age: The diagnostic results, split according to the age group, are given in Table 4.11. Out of the 58 cancers, 44 (75.9%) occurred in women above 50 years of age. Put in

another way, 44 out of the 122 women aged 50 or above (36.1%) had cancer. This proves that age is an important risk factor for the incidence of cancer.

Table 4.11 US dataset: Age distribution and respective diagnoses

Age range	No. of patients
Between 19 and 30 (both ages inclusive)	17
Benign	17
Malignant	0
Between 31 and 40 (both ages inclusive)	35
Benign	32
Malignant	3
Between 41 and 49 (both ages inclusive)	67
Benign	56
Malignant	11
Greater than 50	122
Benign	78
Malignant	44

The other risk factors such as previous breast biopsy, usage of OCP and HRT, smoking or drinking habits, and race, were either not recorded in the study or did not demonstrate significant effect on cancer incidence.

4.2.3 US dataset: Study of effect of lesion parameters on cancer incidence

This section presents the analysis of the effect of lesion parameters on the cancer prevalence among patients.

Lesion location: Table 4.12 presents the lesion distribution at the various regions of the breast. The table also lists the location of malignant lesions. Similar to the TTSH dataset, 48.1% of the lesions appeared in the upper outer quadrant of both breasts and 20.7% in the upper inner quadrant. Moreover, 68.9% of the malignant lesions belonged to upper outer, upper inner, or subareolar regions.

Lesion palpability nature: Among the 241 lesions that were examined, 148 were palpable and 93 were not palpable. Out of the 58 malignant lesions, 41 (70.7%) were palpable.

Lesion size: This parameter was removed from analysis due to the large number of missing values.

Statistical analysis of the BDS datasets

Table 4.12 US dataset: Location of lesions

	Location	No. of lesions	Percentage (%)	No. of malignant lesions	Malignant lesion percentage (%)
Upper outer quadrant	Right	55	48.1	13	43.1
	Left	61		12	
Upper inner quadrant	Right	29	20.7	10	24.1
	Left	21		4	
Lower outer quadrant	Right	16	17.8	6	24.1
	Left	27		8	
Lower inner quadrant	Right	9	8.7	2	6.9
	Left	12		2	
Other locations	Right 3 o'clock position	0	4.6	0	1.7
	Right 6 o'clock position	0		0	
	Right 9 o'clock position	0		0	
	Right 12 o'clock position	0		0	
	Right subareolar	7		1	
	Left 3 o'clock position	0		0	
	Left 6 o'clock position	0		0	
	Left 9 o'clock position	0		0	
	Left 12 o'clock position	0		0	
	Left subareolar	4		0	
Total		241	100	58	100

4.2.4 US dataset: Results of diagnostic tests

Table 4.13 presents the different possible combination of diagnostic tests and the number of patients who took them. All the women were offered biopsy. Unlike the TTSH dataset, the CBE LOS was available in the US dataset.

Table 4.13 US dataset: Patient distribution in terms of diagnostic tests undertaken

Combination of diagnostic tests	No. of patients
Mammography, Ultrasound, BDS, CBE	108
Mammography, BDS, CBE	105
Ultrasound, BDS, CBE	10
Mammography, Ultrasound, BDS	0
Mammography, BDS	1
Ultrasound, BDS	0
CBE, BDS	17

CBE results:

Out of the 241 cases, only one woman was not offered clinical breast examination. Since the patient’s biopsy result indicated cancer and the MG LOS value was 3 (contradictory), the CBE LOS of 4 was assigned to this case. A low sensitivity of 29.31%, high specificity of 98.81%, and an accuracy of 81.33% were registered. The contingency table is presented in Table 4.14.

Table 4.14 US dataset: CBE results (N=241)

	Disease status	
CBE result	Present	Absent
Positive	True positive - 17	False positive - 4
Negative	False negative - 41	True negative - 179
Results		
Sensitivity	29.31%	
Specificity	98.81%	
Accuracy	81.33%	

Ultrasound results:

Out of the 241 cases, 123 women had not taken up ultrasound test. Since there are a lot of cases without ultrasound results, making the LOS assignment for the missing cases (as done in the TTSH dataset) will not provide reliable performance measures. Hence, ultrasound performance measures were not calculated.

Mammography results:

Out of 46 women under the age of 40, 15 women were not offered mammography because of radiation concerns. Overall, out of the 241 cases, 27 women had not taken up mammography test. For these cases, MG LOS value was assigned based on the following conditions:

- If class was malignant and CBE LOS was 4 or 5, MG LOS = CBE LOS
- If class was benign and CBE LOS was 1,2, or 3, MG LOS = CBE LOS
- If class was malignant and CBE LOS was 1,2, or 3, MG LOS = 4
- If class was benign and CBE LOS was 4 or 5, MG LOS = 3

A sensitivity of 43.10%, specificity of 92.90%, and an accuracy of 80.91% were registered. The contingency table is presented in Table 4.15.

Statistical analysis of the BDS datasets

Table 4.15 US dataset: Mammography results (N=241)

	Disease status	
Mammography result	Present	Absent
Positive	True positive - 25	False positive - 13
Negative	False negative - 33	True negative - 170
Results		
Sensitivity	43.10%	
Specificity	92.90%	
Accuracy	80.91%	

BDS results:

The BDS test recorded a sensitivity of 89.66%, specificity of 54.64%, and an accuracy of 63.07%. The contingency table is presented in Table 4.16.

Table 4.16 US dataset: BDS results (N=241)

	Disease status	
BDS result	Present	Absent
Positive	True positive - 52	False positive - 83
Negative	False negative - 6	True negative - 100
Results		
Sensitivity	89.66%	
Specificity	54.64%	
Accuracy	63.07%	

ROC results:

The ROC curves of BDS, mammography, and CBE are illustrated in Figure 4.2. The thresholds that gave the best sensitivity and specificity values and the AUC of these tests are presented in Table 4.17. As seen from Table 4.17 and Figure 4.2, all the three diagnostic tests have clinically acceptable significant AUC values. The low sensitivity of mammography and CBE may be attributed to the lesser number of malignant samples available in this study and hence, even one misclassification would decrease the sensitivity significantly.

Figure 4.2 US dataset: Comparison of ROC curves of CBE, mammography, and BDS tests

Table 4.17 US dataset: Summary of ROC results of CBE, mammography, and BDS tests

ROC RESULTS (Sn: Sensitivity; Sp: Specificity; Th: Threshold)								
CBE N =241			Mammography N =241			BDS N=241		
Th	Sn (%)	Sp (%)	Th	Sn (%)	Sp (%)	Th	Sn (%)	Sp (%)
>=1	100	0	>=1	100	0	>=1	100	0
>1	84.48	24.59	>1	91.38	20.77	>1	91.38	20.77
>2 *	79.31	52.46	>2	91.38	44.26	>2	91.38	28.42
>3	29.31	97.81	>3 *	43.1	92.9	>3 *	89.66	54.64
>4	1.72	100	>4	0	99.45	>4	37.93	94.54
>5	0	100	>5	0	100	>5	0	100
Area under the curve		0.705	Area under the curve		0.755	Area under the curve		0.771

4.2.5 US dataset: Biopsy reduction

In this dataset, there were 205 biopsied cases with a prior BDS LOS of 2 or 3. Among the 205 cases, 169 had a benign biopsy result. In 100 of these cases, the BDS test resulted in Post-BDS LOS of either 2 or 3, i.e. BDS has a 48.78% (100/205) accuracy level in predicting unwanted biopsies.

4.3 Discussion on results obtained using both datasets

In this chapter, the results of the analyses conducted to understand the nature of the two BDS datasets were presented. Also, the performances of the diagnostic tests such as CBE, mammography, ultrasound, and BDS were quantified using sensitivity, specificity, accuracy, and AUC measures. Based on these results, the comparison table (Table 4.18) was drawn to clearly present the similarities and dissimilarities in the results of the two datasets.

Key conclusions from Table 4.18:

The following are the key conclusions that can be drawn from Table 4.18.

- In both the datasets, around 75% of the total malignancies occurred in women over the age of 50 and at least 60% of the total malignancies appeared in post-menopausal women. Around 70% of the malignant lesions originated in the upper quadrants of the breast. Also, at least 70% of the malignant lesions were palpable. This reiterates the fact that age and menopausal status are critical risk factors for breast cancer, and that palpable lesions located at the upper quadrants have a higher probability of being malignant.
- The TTSH BDS test resulted in higher performance measures compared to those obtained from the US dataset BDS test. The probable reason for this significant improvement is the difference in the interpretation maps and BDS index calculation algorithms used in both the devices.
- There is a lot of scope for accuracy improvement in the BDS US test. Classifiers may be a viable option for improving the diagnostic accuracy.
- It was mentioned earlier that the prior LOS values are determined by the doctor by using the mammography LOS, ultrasound LOS, and/ or CBE LOS. This prior LOS is used in the BDS interpretation map (Table 2.3 or 2.4) to determine the Post-BDS LOS. It is to be noted that the respective LOS values are used to calculate the performance measures of a test (i.e., Mammography LOS is used to determine mammography accuracy, ultrasound LOS is used to determine ultrasound accuracy, and so on). Thus, the Post-BDS LOS is dependent on the prior LOS, which in turn is dependent on the LOS of other tests. This dependence leads to the reflection of the other test results on the BDS test results. This can be inferred from the data presented in Table 4.18 - the TTSH BDS test results are quite similar to the

corresponding ultrasound results. The US BDS test results, on the other hand, are different from the corresponding mammography/ CBE results due to the different interpretation map used. Nevertheless, an important point to note here is that the results of these other tests are subjective, and tend to vary based on the capabilities of different doctors involved. Hence, in spite of the use of a quantitative BDS index for predictions, due to the dependence highlighted above, the BDS interpretation methodology can be considered as slightly subjective. If there were any way to eliminate the Post-BDS LOS, and use only the BDS index and/or raw sensor electropotential values for disease detection, the BDS test can be developed into a more objective adjunct modality. Classification tools can be exploited for this purpose. The next chapter presents the analysis in this angle.

- In the case of the US dataset, with the help of the BDS results it is possible to reduce the unwanted biopsies by around 49%; whereas a 93% possible biopsy reduction is seemingly possible in the TTSH BDS test. This improvement may again be attributed to the modified device configuration used in the TTSH BDS device. Overall, based on the significant biopsy reduction percentages, it is evident that the BDS device does have the capability of giving a good second opinion, which is the key necessity of any adjunct modality.

Statistical analysis of the BDS datasets

Table 4.18 Comparison of TTSH and US datasets

	TTSH DATASET	US DATASET
Total number of samples	182	241
Normal	33	Not available
Benign	96	183
Malignant	53	58
CONFIGURATION OF THE BDS DEVICE USED (Refer section 2.4.6.5)		
Interpretation map used	<i>Map 2</i>	<i>Map 1</i>
Sensor arrangement used	<i>SA 2</i>	<i>SA 2</i>
BDS index calculation using	<i>Algo 2</i>	<i>Algo 1</i>
Pre-processing steps	Missing value handling; No outliers	Missing values and outlier handling
EFFECT OF KEY RISK FACTORS		
Age (% of cancers in women aged 50 and above)	(40/53 cancers) : 75.5%	(44/58 cancers) : 75.9%
Post-menopausal status (% of cancers in post-menopausal women)	(33/53 cancers) : 62.3%	(44/58 cancers) : 75.9%
EFFECT OF LESION PARAMETERS		
Lesion location (% of cancers in upper quadrant)	81%	68.9%
Lesion palpability status (% of palpable malignant lesions)	(47/53 cancers) : 88.7%	(41/58 cancers) : 70.7%
Lesion size	Malignant lesions were at least 6 mm along the longest diameter	Feature removed due to large number of missing samples
DIAGNOSTIC PERFORMANCE MEASURES		
BDS		
Sensitivity (%)	96.23	89.66
Specificity (%)	93.80	54.64
Accuracy (%)	94.51	63.07
AUC	0.972	0.771
Mammography		
Sensitivity (%)	90.57	43.10
Specificity (%)	98.45	92.90
Accuracy (%)	96.15	80.91
AUC	0.945	0.755
Ultrasound		
Sensitivity (%)	96.23	Many missing samples
Specificity (%)	94.57	Many missing samples
Accuracy (%)	95.05	Many missing samples
AUC	0.975	Many missing samples
Clinical Breast Examination (CBE)		
Sensitivity (%)	CBE LOS not available	29.31
Specificity (%)	CBE LOS not available	98.81
Accuracy (%)	CBE LOS not available	81.33
AUC	CBE LOS not available	0.705
POSSIBLE BIOPSY REDUCTIONS		
No. of biopsied cases with prior LOS of 2 or 3 - (i)	44	205
No. of benign biopsies among (i) - (ii)	42	169
No. of cases among (ii) that had a Post-BDS LOS of 2 or 3 - (iii)	41	100
% biopsy reduction - (iii/i)*100	93.18	48.78

CHAPTER 5: EVALUATION OF THE DATA MINING FRAMEWORK

In the previous chapter, it was concluded that various classifiers could be employed to improve the prediction accuracy of the BDS test and also to reduce the subjective nature of the current BDS interpretation methodology. Classification is only a part of the whole data mining process. Before a dataset is subjected to any classification technique, it has to be cleaned to remove outliers and irrelevant or redundant features. In view of that, a generalized data mining framework encompassing all the key steps was developed. The two BDS datasets were studied using this framework. The results and key findings are presented in this chapter.

5.1 Proposed data mining framework

The general framework of data mining consists of the following key steps: (1) Data pre-processing, wherein the missing values and outliers are handled and the data is normalized, (2) Feature selection, wherein good feature subsets are selected using filter and wrapper techniques, (3) Classifier development, wherein the selected feature subsets are used to build classifiers, (4) Performance evaluation, wherein the built classifiers are evaluated using test data to determine their accuracy in classifying the data. The techniques employed in each of these steps have been described in section 2.7. The proposed framework is illustrated in Figure 5.1. In the figure, “*Dataset(FeatureSet)*” implies that the classifier was developed or evaluated using the reduced dataset formed by extracting only the selected *FeatureSet* from the *Dataset*. *FFS* stands for Filter Feature Subset and *WFS* stands for Wrapper Feature Subset.

According to the framework, missing data, and the univariate and multivariate outliers present in the analyzed dataset are first handled. Subsequently, all the available features are encoded, and normalized. These pre-processed features are then used in filter and wrapper feature selection algorithms to select feature subsets. In this work, ten-fold stratified cross validation technique was used for feature selection. The selected features are then used to form a reduced dataset of the original dataset. This reduced dataset is used to build and test classifiers using ten-fold stratified cross validation. That is, the whole dataset is split into ten parts such that each part contains

approximately the same proportion of class samples as the original dataset. Nine parts of the data (training/learning set) are used for classifier development and the built classifier is evaluated using the remaining one part (test set). This procedure is repeated ten times using a different part for testing in each case. The ten test classification performance metrics are then averaged, and the average test performance is declared as the estimate of the true generalization performance.

Figure 5.1 Proposed data mining framework

5.2 Pre-processing

A summary of the overall features used from each dataset, details of how the features were encoded before useful features were extracted, is given in this section.

5.2.1 TTSH dataset

In this dataset, the size parameter was missing in three cases, and hence, the class average of size was used to counter that. No outliers were present. The following features were used in the data mining: Menarche age, Menopausal status, Pregnancy history, Family history, Lesion palpability status, Lesion location, Lesion size, Age, BDS index, Post-BDS LOS, and Class. It is easier for analysis if the string variables

are represented as nominal/numeric variables. Hence, the following encoding was done. Read as

Feature - Old representation: New representation.

- Menopausal status – Pre: 0; Post: 1
- Pregnancy history – Previous pregnancy (no risk):0; Never pregnant (has risk): 1
- Family History – No history: 0; Has history: 1
- Lesion palpability status – Not palpable: 0; Palpable: 1
- Lesion location – Since lesion location details will not be available for the normal samples, the following encoding was used for uniformity.

RLO, RLI, right 6 o'clock position, LLO, LLI, left 6 o'clock position, Normal cases: 0

All other locations: 1

Where (R–Right; L–Left; U–Upper; L–Lower; I–Inner; O–Outer; SA–Sub Areolar)

- Class/Biopsy – Benign, Normal: 0; Malignant: 1

The remaining parameters were used as such, as they were numerical quantities. The total number of features in the TTSH dataset, including the class attribute, amounts to 11. Once the features were encoded, they were normalized using min-max normalization.

5.2.2 US dataset

The information about how the missing data were handled and outliers were removed in this dataset was given in section 4.2.1. The following features were used in the data mining: Menopausal status, Pregnancy history, Family history, Lesion palpability status, Lesion location, Age, 16 electropotential values - SC, SU, SL, SI, SO, SV, SH, SD, AC, AU, AL, AI, AO, AV, AH, AD, BDS index, Post-BDS LOS, and Class. The following encoding was done. Read as

Feature - Old representation: New representation.

- Menopausal status – Pre: 0; Post: 1
- Pregnancy history – Previous pregnancy (no risk):0; Never pregnant (has risk): 1
- Family History – No history: 0; Has history: 1
- Lesion palpability status – Not palpable: 0; Palpable: 1
- Lesion location – RUO: 1;RUI: 2; RLO: 3; RLI: 4; RSA: 5; LUO: 6; LUI: 7; LLO: 8; LLI: 9; LSA: 10

- Class/Biopsy – Benign: 0; Malignant: 1

The remaining parameters were numerical quantities. Sometimes, linear combinations of features aid in better classification than the original features. Hence, ten additional features were derived using the 16 raw electropotentials. These electropotential differentials (EPD), which are listed in Table 5.1, were derived so as to reflect the mean electropotentials or electropotential differentials within the symptomatic breast or the differentials between the symptomatic and asymptomatic breasts. Since studies (Faupel et al., 1997) have demonstrated the usefulness of a few such differentials (EPD 3 and EPD 6) in differentiating malignant and benign lesions, only differentials similar to these EPDs were derived in this work. Thus, including the class attribute, the number of features amounts to 35. All the features were normalized using min-max normalization.

Table 5.1 Derived electropotential based features

Additional derived features; EPD - Electropotential Differentials
EPD 1 : Mean(values from all symptomatic sensors SC, SU, SL, SI, SO, SV, SH, SD)
EPD 2 : Mean(values from symptomatic sensors placed in the examined lesion's quadrant: SC,SU,SI,SL,SO)
EPD 3 : SC-SV
EPD 4 : Within breast differential : 3SC-SD-SH-SV
EPD 5 : Within breast differential : 4SC-SI-SU-SO-SL
EPD 6 : Between breasts differential : EPD 1 - Mean(values from all asymptomatic sensors AC, AU, AL, AI, AO, AV, AH, AD)
EPD 7 : Between breasts differential: EPD 2 - Mean(values from asymptomatic sensors AC,AU,AI,AL,AO)
EPD 8 : EPD 4 - (3AC-AD-AH-AV)
EPD 9 : EPD 5 - (4AC-AI-AU-AO-AL)
EPD 10 : EPD 3 - (AC-AV)

5.3 Feature selection

As proposed in the framework, feature selection was done using wrapper and filter feature selection techniques. In some methodologies, filter techniques are used first, followed by wrapper techniques to select the best feature subset. Selecting a feature subset using one technique first, and then filtering out the best from this selected subset using another technique might result in losing what would have been some important features. Therefore, in this work, both these techniques were applied independently to the dataset to get the respective feature sets which were then used to build and evaluate various classifiers. Finally, the best feature subset-classifier

combination was chosen for the studied dataset based on the performance measures such as sensitivity, specificity, accuracy, and AUC values.

5.3.1 TTSB dataset

MedCalc was used to calculate the non-parametric measure of correlation called *Spearman's rank correlation (rho)* between each of the features and the class. Among all the features, significant correlation coefficient values were observed for menopausal status (0.322), age (0.386), lesion size (0.451), lesion palpability status (0.402), BDS index (0.397), and Post-BDS LOS (0.797). The p-values in all cases were less than 0.0001. Hence, they appear to be highly informative features.

To confirm the goodness of these features, filter techniques described in section 2.7.5, namely, Correlation-based Feature Selection (CFS), Chi-Squared method (CS), Information Gain (IG), Gain Ratio (GR), Symmetrical Uncertainty (SymU), and ReliefF methods, were used. The software WEKA was used for feature selection. Ten-fold cross validation was used on the whole dataset, and the features were finally ranked using the average of scores obtained in each fold. The top three features (25%) were selected to form each feature subset. It was observed that a similar subset of {Post-BDS LOS, lesion size, lesion palpability} was selected by CS, IG, GR, and SymU methods. Hence, these three features selected by the Chi-Squared method can be considered as the representative best subset *FFSI* produced by the four filters (CS, IG, GR, and SymU). Since the ranks varied significantly for CFS and ReliefF methods, the top three features selected by these methods are listed separately in Table 5.2. Table 5.2 also lists *FFSI*.

The wrapper feature subset (WFS) was selected using an LDA classifier that was implemented using Matlab. The LDA classifier was chosen because it is a computationally less intensive classifier that is fast and efficient for repetitive evaluation by the wrapper method using various feature sets. The search strategy employed was nondeterministic strategy and the search direction was random. The ranks of the features in every fold were considered to form the final subset. The resulting *WFS* subset had the following features: Post-BDS LOS, lesion size, and palpability status (also listed in Table 5.2).

Evaluation of the data mining framework

Table 5.2 TTSH dataset: Features selected using filter and wrapper methods

Chi-Squared method <i>FFS1</i>	CFS method <i>FFS2</i>	ReliefF method <i>FFS3</i>	Wrapper LDS method <i>WFS</i>
Post-BDS LOS Size Palpability	Post-BDS LOS BDS index Size	Post-BDS LOS Palpability Location	Post-BDS LOS Size Palpability

5.3.2 US dataset

The filter and wrapper techniques used on the TTSH dataset were also used for the US dataset. Again, ten-fold cross validation was used. The top eight features (25%) were selected to form each feature subset. The features in Table 5.3 are listed (top eight based on average scores) according to their chi-square-statistic. However, it can be observed that the ranks of these features according to IG, GR, and SymU methods do not vary a lot. Hence, the eight features selected by the Chi-Squared method can be considered as the representative best subset *FFS1* produced by the four filters (CS, IG, GR, and SymU). Also, the features selected by the CFS method are similar to those in *FFS1*. Table 5.4 lists the feature subsets *FFS1*, *FFS2*, *FFS3*, and *WFS* that were selected for the US dataset.

Table 5.3 US dataset: Top five features according to Chi-Squared method and the ranks of these features according to GR, IG, and SymU filters.

Feature description	Chi-Squared evaluator/ Ranker search		Gain Ratio evaluator/ Ranker search		Information Gain evaluator/ Ranker search		Symmetrical Uncertainty evaluator/ Ranker search	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank
EPD8	50.495	1	0.291	1	0.142	2	0.221	1
Post-BDS-LOS	50.877	2	0.116	6	0.171	1	0.151	4
EPD4	43.123	3	0.282	2	0.121	3	0.197	2
EPD5	40.181	4	0.257	3	0.112	4	0.181	5
EPD7	34.463	5	0.249	4	0.097	5	0.162	3
EPD3	28.385	6	0.165	5	0.08	6	0.124	6
SH	21.171	7	0.072	8	0.071	7	0.08	7
SO	18.587	8	0.073	7	0.056	8	0.076	8

Evaluation of the data mining framework

Table 5.4 US dataset: Features selected using filter and wrapper methods

Chi-Squared method <i>FFS1</i>	CFS method <i>FFS2</i>	ReliefF method <i>FFS3</i>	Wrapper LDS method <i>WFS</i>
EPD8	EPD4	Palpability	EPD2
Post-BDS-LOS	EPD5	Menopause Status	EPD1
EPD4	EPD7	Post-BDS LOS	EPD4
EPD5	EPD8	EPD4	EPD5
EPD7	Post-BDS LOS	EPD8	SC
EPD3	EPD3	EPD5	SV
SH	SH	EPD3	SO
SO	SO	SH	EPD7

5.4 Classification results

In this work, the reduced learning datasets formed using the filter feature subsets *FFS1*, *FFS2*, and *FFS3* were used to build the following classifiers - BPNN, PNN, LDA, SVM, and FIS. Since wrapper feature sets should be used with their respective classifiers, the reduced learning dataset formed using wrapper feature subset *WFS* was used to build a LDA classifier (LDA/*WFS* classifier). Matlab was used for classifier development and evaluation. As done for feature selection, stratified ten-fold cross validation technique was employed for building and evaluating these classifiers. This resampling technique is widely used for datasets with less number of samples, like the datasets used in this work, to get robust performance results. In each fold, a threshold of 0.5 was used to round off the classifier output (predicted class) to belong to either class 0 or 1. The performance measures were calculated by comparing the predicted class and the actual class attribute of the data.

5.4.1 TTSH dataset

The TTSH BDS test resulted in a sensitivity of 96.23%, specificity of 93.80%, accuracy of 94.51%, and AUC of 0.97. The classification results obtained using the filter subsets are presented in Table 5.5. In the case of BPNN, a two layer network was designed with the first layer having five tansig neurons (empirically chosen) and the second layer with one tansig neuron. Levenberg-Marquardt Back Propagation training algorithm was used because it is the fastest. Gradient descent with momentum weight and bias learning function was used for learning. The PNN was built using a spread of 0.2. For SVM, a Radial Basis Function (RBF) kernel function with a scaling factor of 2 (empirically chosen) was used to map the training data into the kernel space. In the

Evaluation of the data mining framework

case of FIS, the radius was empirically chosen as 3. It is evident from Table 5.5 (highlighted cells) that irrespective of the feature subset used, LDA and FIS resulted in similar performance measures that are equal to the measures recorded by the TTSH BDS clinical test. Also, the FFS2 dataset always provided the same results, irrespective of the classifier used. The LDA/WFS classifier also recorded similar high sensitivity of 96.23%, specificity of 93.80%, accuracy of 94.51%, and AUC of 0.97.

Table 5.5 TTSH dataset: Performance measures of the five classifiers

Classifier	Sensitivity (%)			Specificity (%)			Accuracy (%)			AUC		
	FFS1	FFS2	FFS3	FFS1	FFS2	FFS3	FFS1	FFS2	FFS3	FFS1	FFS2	FFS3
BPNN	92.45	96.23	90.57	95.35	93.80	93.80	94.51	94.51	92.86	0.97	0.97	0.95
PNN	88.68	96.23	86.79	94.57	93.80	94.57	92.86	94.51	92.31	0.92	0.97	0.91
LDA	96.23	96.23	96.23	93.80	93.80	93.80	94.51	94.51	94.51	0.97	0.97	0.97
SVM	96.23	96.23	94.34	93.02	93.80	93.80	93.96	94.51	93.96	0.95	0.97	0.94
FIS	96.23	96.23	96.23	93.80	93.80	93.80	94.51	94.51	94.51	0.97	0.97	0.97

5.4.2 US dataset

The US BDS test recorded a sensitivity of 89.6%, specificity of 54.6%, accuracy of 63.1%, and AUC of 0.77. The classification results obtained using the filter subsets are presented in Table 5.6. In the case of BPNN, a two layer network was designed with the first layer having eight tansig neurons (empirically chosen) and the second layer with one tansig neuron. Levenberg-Marquardt Back Propagation training algorithm was used because it is the fastest. Gradient descent with momentum weight and bias learning function was used for learning. The PNN was built using a spread of 0.09. For SVM, a RBF kernel function with a scaling factor of 2 (empirically chosen) was used to map the training data into the kernel space. In the case of FIS, the radius was empirically chosen as 0.5. It can be observed from Table 5.6 that the results for *FFS1* and *FFS2* are the same, as they have the same features. Even though PNN, SVM, and FIS demonstrate good specificity values (higher than 95%), the corresponding sensitivity values were low. The highest AUC value of 0.84 was recorded by BPNN for the *FFS3* feature subset. Irrespective of the feature subsets used, every classifier resulted in almost equal performance measures for all the feature subsets. However, the LDA/WFS classifier recorded a high sensitivity of 93.10%, specificity of 98.36%, accuracy of 97.10%, and AUC of 0.96.

Evaluation of the data mining framework

Table 5.6 US dataset: Performance measures of the five classifiers

Classifier	Sensitivity (%)		Specificity (%)		Accuracy (%)		AUC	
	FFS1/ FFS2	FFS3	FFS1/ FFS2	FFS3	FFS1/ FFS2	FFS3	FFS1/ FFS2	FFS3
BPNN	74.14	75.86	91.80	92.90	87.55	88.80	0.83	0.84
PNN	46.55	44.83	96.72	95.08	84.65	82.99	0.71	0.70
LDA	75.86	70.69	77.05	80.87	76.76	78.42	0.76	0.75
SVM	63.79	67.24	97.27	96.17	89.21	89.21	0.80	0.82
FIS	62.07	37.93	95.63	80.87	87.55	79.44	0.79	0.72

5.5 Discussion on results obtained using both datasets

The results obtained by using the data mining techniques on the two BDS tests were presented in this chapter. The following are the key observations made based on the results:

TTSH dataset:

- For the TTSH dataset, it was found that all the classifiers resulted in similar performance measures for *FFS2*, which were equal to those obtained using the TTSH BDS test. LDA/WFS classifier also presented similar measures. The TTSH BDS clinical test accuracy was calculated using the Post-BDS LOS feature. The sole aim of using classifiers was to verify if additional features like demographic parameters and lesion details would improve the accuracy. As is evident from the classification results, this is not the case for this dataset. The reason can be attributed to the fact that the Post-BDS LOS is an overpowering feature with a relatively high correlation with the final class ($\rho > 0.797$). The other features ($\rho < 0.5$) did not provide any more discriminatory information than the Post-BDS LOS. To verify this, all the classifiers were evaluated with just the Post-BDS LOS as a single feature. As was expected, similar performance measures, as those obtained with the initial three features, were obtained. This shows that adding the other two features on top of Post-BDS LOS has not had any effect on the classification results. Thus, for this dataset, classifiers would not have any improved effect unless other stronger discriminatory features like, for example, the raw electropotential values measured using the sensors, are used. This observation presented the need to study a dataset with these electropotential values, and hence, the US dataset was acquired and analyzed in this work.
- For completion sake, the TTSH dataset was analyzed with only 96 benign and 53 malignant samples, omitting the normal cases (a total of 149 samples). Again, Post-

BDS LOS was found to be strongly influencing the classification results. If only 149 cases are considered, the BDS clinical test sensitivity, specificity, and accuracy values were 96.23%, 91.67%, and 93.29%, respectively. Similar measures were obtained by using all the classifiers.

US dataset:

- In the case of the US dataset, BDS clinical test demonstrated a good sensitivity of 89.6%. However, the specificity (54.6%) and accuracy (63.1%) were low. The AUC was also only 0.771. In the case of the filter results, it was observed that even though the obtained accuracies were over 75% for most classifiers, the sensitivities were low and the specificities were high. Generally, for a good classifier, it would be ideal to have a balance between these two measures and have high values for the same. In this regard, only BPNN and LDA were acceptable. BPNN had good AUC values of around 0.83, and LDA around 0.76, without any drastic imbalance between the sensitivity and specificity values.
- The LDA/WFS classifier/feature set combination recorded a significantly high sensitivity of 93.10%, specificity of 98.36%, accuracy of 97.10%, and AUC of 0.96. To summarize, both the filter and wrapper methods produced feature subsets which resulted in better classification accuracies (at least 75% for the filter and wrapper subsets) when compared to that given by the original BDS device (63.1%).
- The higher accuracy of LDA wrapper technique compared to that of LDA filter technique reconfirms that wrapper methods are indeed better for feature selection, if computational time is not a key concern.
- It can be observed that none of the feature subsets had any of the electropotentials recorded from the asymptomatic side as features. They either selected symptomatic sensor values or the difference between symptomatic and asymptomatic electropotentials, thereby, emphasizing the fact that the surface electropotentials are reliable indicators of the presence or absence of cancer.
- It can be seen that *FFS1*, *FFS2*, and *FFS3* feature subsets had the Post-BDS LOS as one of the selected features, whereas *WFS* did not have this feature. However, among all the feature subsets, *WFS* resulted in the highest sensitivity, specificity, accuracy, and AUC. The most important observation here is that even without using the Post-BDS LOS, *LDA/WFS* was able to classify with clinically acceptable

sensitivity and specificity results. As discussed earlier, the Post-BDS LOS is calculated using the prior-LOS and BDS index. This prior LOS is determined by the ultrasound/mammogram/physical examination results of the patient. Thus, it is evident that the Post-BDS LOS is dependent on the results of the other tests for its determination. These other tests are relatively subjective, as their LOS values are decided by the doctor. This dependence indirectly makes the BDS test slightly subjective. Since LDA/WFS presented good results without using this Post-BDS LOS, it can be concluded that by using this classification scheme, there is a possibility that the BDS result (presence/absence of cancer) can be predicted using only the original and derived electropotentials, i.e., the determination of the prior LOS from the patient's ultrasound/mammogram/physical examination results and, hence, the Post-BDS LOS is not necessary. Thus, by using this classification scheme, BDS can be regarded as an independent diagnostic device, as its results can be obtained without relying on ultrasound and mammogram results. However, BDS cannot replace established primary tests like mammography and ultrasound, as it is still categorized as an adjunct device that is used to provide a second opinion after the patient is examined using mammography and ultrasound. This conclusion could not be validated using the improved TTSH BDS device, because of the unavailability of the 16 raw electropotential values in the TTSH dataset. Nevertheless, the device improvement was only reflected in the calculation of the BDS index and Post-BDS LOS, and both these features were not used in obtaining the above conclusion. Therefore, there is a high possibility that the conclusion holds true for the TTSH dataset too.

Any classification problem can be approached using the proposed framework first. Consequently, based on the results, the analysis can be further carried out using more complicated classifiers, if deemed necessary. The main advantage of this technique is that many feature subset-classifier combinations may be evaluated on the dataset, leading to an improved probability of finding a more appropriate feature set and corresponding classifier for the analyzed dataset. Since k-fold resampling technique is employed, it can be used to get robust classification performance measures for any dataset irrespective of the sample size. Also, in the case of the US dataset, if the filter techniques were employed first to choose a subset of features, followed by wrapper techniques to further select a good subset from the filter subsets, there is a possibility

that the wrapper subset *WFS* would have not been identified at all. Thus, because of the use of both these techniques independently on the dataset, a potentially good *WFS* subset was identified.

In the previous chapter, two goals were set - First, to use the classification tools to improve the diagnostic accuracy of the BDS datasets; second, to use the classification tools to help in reducing the mild subjectivity associated with the BDS prediction methodology. Accordingly, the classifiers were evaluated and both the goals were met. The one limitation that was observed was that some of the classifiers used were not transparent, i.e. the way in which the classifiers determine the final diagnostic result is not comprehensible by the end users (for e.g., BPNN). In some classifiers, the interpretation is simple and easily understood. For example in the case of the LDA classifier, the linear combination of features that separates the classes can be obtained. However, when the sample size increases, or when both the classes have equal number of samples unlike the US dataset used in this work, there is a possibility that the classes might not have the good linear separability observed now. Therefore, it would be worthy to develop a more transparent classifier that would not be affected by class imbalance and that would also provide comparable or better classification accuracy. The next chapter presents the algorithm of one such classifier that was developed in this work.

CHAPTER 6: WRUBAC- DEVELOPMENT AND EVALUATION

WRUBAC stands for Weighted Rule Updation Based Associative Classification. It is a type of associative classifier whose algorithm incorporates the traditional three steps, namely, rule generation, rule pruning, and classification. Apart from these steps, the algorithm also has a rule update process wherein the weight of each rule is updated based on the training dataset.

6.1 WRUBAC algorithm

The steps used in WRUBAC algorithm are presented in detail in this section.

1. Data pre-processing

- 1.1 Analyze the dataset and handle any missing attribute values.
- 1.2 Scan the dataset for outliers and remove them with careful discretion.
- 1.3 Apply stratified sampling to randomly select 70% of the dataset for training (for feature selection and associative classifier development) and preserve the remaining 30% for testing.
- 1.4 Load the training data into WEKA.
- 1.5 Apply supervised (class-based) discretization algorithm, which is based on the methodology proposed in Fayyad and Irani (1993), and discretize every continuous variable. Make a note of the discretization intervals.
- 1.6 Apply filter and wrapper feature selection techniques (as highlighted in the proposed framework) and make a note of the feature subsets selected by each technique.
- 1.7 Write a customized discretization function for the studied dataset in Matlab. The function, when called from a main program, should apply the discretization intervals (obtained earlier from WEKA) to the studied dataset. In this function, the values falling within the first discretization interval are assigned the code 1, and so on. Moreover, the function should also append the feature id to each code, i.e., if the third feature has three discretization codes, then it has to be coded as 31, or 32, or 33 depending on the code. This method has to be employed to identify the features in the extracted rules. This function

is written separately for ease of use on the test data or any other new data that may be tested at later times.

- 1.8 Load the training and test datasets into Matlab and discretize them using the function written earlier.
- 1.9 Study one feature subset at a time. Hence, form reduced train dataset and reduced test dataset by retaining only the selected features in each of the datasets (**Pruning level 1**). Note, for the training dataset, the class feature is present in the dataset as it is needed for building WRUBAC (it is a supervised learning classifier). However, the test dataset has only the features, and the desired class values are stored separately.

2. Rule generation: training dataset to be used

- 2.1 Determine the number of classes (*NoOfClasses*) in the train dataset. Let $i = 0$, $MaxAntecedantLength = \text{Number of features in the dataset}$, $RuleLength = MaxAntecedantLength + 1$;
- 2.2 $i = i + 1$.
- 2.3 Select the samples belonging to class i , and extract association rules from these samples using Apriori algorithm (Matlab version implemented in ARMADA). Use minimum support of 10% (**Pruning level 2**). Each rule in the rule set for this class will have the following details:
 - a. LHS - the antecedents of the rule; the maximum number of items in the LHS will be equal to the total number of features in the train set. For example, assume the dataset has three features F1, F2, and F3. The rule set can have one-itemsets of the form (F1: value), (F2: value), (F3: value); two-itemsets of the form {(F1: value), (F2: value)}, {(F2: value), (F3: value)}; or three-itemsets of the form {(F1: value), (F2: value), (F3: value)}.
 - b. RHS - the consequent of the rule. ARMADA extracts both classification and association rules. Note: ARMADA does not extract rules if the class label is 0 in the dataset. Care should be taken to label the classes as 1 or greater than 1.
 - c. Support in number.
 - d. Confidence percentage.

- 2.4 **Pruning level 3**: Discard the rules with a confidence value less than 50%. It has been found that the 50% confidence threshold is the maximum threshold that can be used to prune the rule set without affecting accuracy (Chiusano et al., 2009).
- 2.5 **Pruning level 4**: Discard the rules which do not have the class attribute on the RHS. i.e., retain only the classification rules.
- 2.6 For each rule in the pruned rule set, assign support in percentage and a rule weight defined as

$$RuleWeight = \frac{Support(\%) * Confidence(\%)}{10000}$$

Generally, support and confidence are the two most important parameters in association rule mining. Some rules might have high confidence, but low support. By using their product as weight, the effect of each of these parameters is balanced.

- 2.7 To sum up, each class rule will have the following details: LHS, RHS, support in number, confidence in percentage, support in percentage, and *RuleWeight*. Repeat steps 2.2 to 2.6 for each of the remaining classes.
- 2.8 Combine all the class rules into one rule base, called *Pruned12AllClassRules*

3. Rule pruning: *Pruned12AllClassRules* to be used

At this stage, the rule base will have both unique and similar classification rules. Similar rules are the rules that have the same LHS, but different class labels. They are otherwise known as conflicting rules. The role of this pruning step in WRUBAC is to keep unique rules, select similar rules that are to be retained, and discard the rest. This pruning step, a novel one in this work, is named as the *conflicting rule weight based pruning* (**Pruning level 5**).

- 3.1 Append a column of 3s to the rule base (This 3 will later be replaced by 1 for rules that are to be removed and by 2 for rules that are to be kept). It acts as a decision flag.
- 3.2 Let $e = 0$; e is the counter for current rule.
- 3.3 $e = e + 1$; *TempClassOfRule* = a zero matrix of size $2 \times NoOfClasses$
- 3.4 For every rule (let it be called as the *comparison rule*) in the rule base that still has the decision flag set to 3,

- a. Compare the current rule with the *comparison rule*.
- b. If there is a match in LHS, store the rule number (index) of the *comparison rule* in $TempClassOfRule(1, ClassLabelValue)$, and its corresponding weight in $TempClassOfRule(2, ClassLabelValue)$. $ClassLabelValue$ is the class label value present in the RHS of the *comparison rule*.
- c. Repeat step 3.4 until all the rules are compared with the current rule.
[For example, assume the current rule 1 belongs to class 1 with weight 0.9, the first matched comparison rule 6 belongs to class 2 with weight 0.8, and the second matched comparison rule 11 belongs to class 3 with weight 0.1. Note, there can only be as many matched rules as the number of classes. In this case, $TempClassOfRule$ will be of the form

$$TempClassOfRule = \begin{bmatrix} 1 & 6 & 11 \\ 0.9 & 0.8 & 0.1 \end{bmatrix}$$

3.5 If the number of zeros in the first row of $TempClassOfRule$ is equal to one less than the number of classes, it indicates that there was only one self-match (current rule matched with itself). Therefore, set the decision flag of the current rule to 2 as the current rule is a unique one.

3.6 If there have been matches, then calculate the following:

- a. Find the maximum rule weight in the second row of $TempClassOfRule$ (e.g., 0.9).
- b. Determine a difference array $MaxDiff$ by subtracting the maximum weight from each of the elements in the second row of $TempClassOfRule$ (e.g. $MaxDiff = [0 \ 0.1 \ 0.8]$).
- c. Convert $MaxDiff$ into a logical array $GTLogicalMat$ by coding any element > 0.5 as 1, 0 otherwise (e.g. $GTLogicalMat = [0 \ 0 \ 1]$).
- d. If $GTLogicalMat$ is all zeros, it indicates that the weight difference between rule with maximum weight and all other matching rules is less than 0.5. This means, all the rules have near equal weights for all the classes they satisfy, and hence, will not be useful in discriminating one class from another. Hence, set the decision flag of all these rules to 1. The rules can be efficiently located in the rule base by using their indices already stored in the first row of $TempClassOfRule$.

- e. If *GTLogicalMat* has one zero, it indicates that the weight difference between the rule with maximum weight and all other matching rules is greater than 0.5. This means, the rule with maximum weight has higher discriminatory power for its class. Hence, change its decision flag to 2 and the decision flag of the other matched rules to 1.
- f. If *GTLogicalMat* has a single "1", it indicates that a rule with a weight difference of greater than 0.5 between itself and the other matched rules exists (as in the illustrated example, where the third rule has the lowest weight, and a *greater than 0.5* weight difference with the other two rules). This means that this low weight rule has the capability of separating its class from the other classes, if assigned a negative weight. Hence, change its decision flag to 2, its weight to negative of current weight, and the decision flag of the other matched rules to 1 (negative rule weighting).
- g. If *GTLogicalMat* has equal number of ones and zeros, none of the rules have better discriminatory power than the rest, and hence, mark them all for deletion.

3.7 Repeat steps 3.3 to 3.6 until all rules are verified.

3.8 Retain/ remove the rules based on their decision flag value.

3.9 Store the pruned rules in *Pruned3Rules*.

4. Rule updation: *Pruned3Rules* to be used

The general concept here is to use the rules in *Pruned3Rules* set to classify the training samples and obtain the predicted class for each sample. The weight of a rule is then updated based on its participation in the number of correctly and number of incorrectly classified samples.

4.1 Classification of the training set using rules in *Pruned3Rules*. For each sample do the following:

- a. For each set of class rules: Add the *RuleWeight* of every rule whose LHS items are present in the current sample to get the cumulative rule weight for that class.
- b. Repeat step (a) for the remaining classes.
- c. Assign the class label that has the maximum cumulative weight as the predicted class for that sample.

- d. Repeat steps (a) to (c) for all the samples in the training set.

4.2 For each sample, do the following:

- a. If the predicted class is equal to the target class, then all the rules having a LHS match with the current sample get their *WtMaintain* counter incremented by 1. In other words, all the rules that classified the sample have their *WtMaintain* counters incremented by 1; i.e. these rules get bonus point for weight maintenance as they correctly classified the sample.
- b. If the predicted class is not equal to the target class,
 - i. The rules in the rule base belonging to the wrongly predicted class get their *WtLoss* counter incremented by 1; i.e. these rules have to undergo weight loss as they wrongly predicted this class.
 - ii. The rules in the rule base belonging to the actual target class get their *WtGain* counter incremented by 1; i.e. these rules have to gain weight in order to correctly assign this class to the sample.
 - iii. The rules in the rule base belonging to any other class get their *WtMaintain* counter incremented by 1; i.e. these rules had the weight not to predict this class, so they need no weight update.

4.3 For each rule, do the following:

- a. Calculate the rule weakness as follows:

$$RuleWeakness = \frac{WtGain + WtLoss}{WtMaintain}$$

- b. Calculate the rule goodness as follows:

$$RuleGoodness = 1 - RuleWeakness$$

4.4 Calculate the class rule goodness *RGMean* by calculating the mean of the rule goodness of all the rules in one class. Repeat for all the other classes.

4.5 Define Rule Goodness Threshold (RGT) as

$$RGT (Class i) = RGMean (Class i) - InputThreshold$$

InputThreshold is a parameter that is to be entered by the user. Prune the class rules whose *RuleGoodness* is less than RGT (**Pruning level 6**). This step is named as *rule goodness threshold based pruning* and is novel in WRUBAC.

WRUBAC - development and evaluation

For example, if the *RGMean* for class 1 is 0.9, and *InputThreshold* is 0.2, then all the rules with *RuleGoodness* less than 0.7 will be pruned. If the RGT for a class rule set had been set as a constant 0.5, and if all the rules in the class had rule goodness less than 0.5, then all rules would have been pruned. Therefore, this parameter has been kept as a variable that can be set by the user after scrutinizing the goodness of rules extracted.

- 4.6 Input an *UpdateConst* (update constant) parameter from the user. This is used to control the percentage of update that is applied to each rule. It can vary from 0 (no update) to any positive or negative integer. This parameter has to be tuned empirically.
- 4.7 Update the weight of each rule using the following formula

$$NewRuleWeight(rule) = RuleWeight(rule) + \frac{(UpdateConst * (WtGain(rule) - WtLoss(rule)))}{NC}$$

where NC (Normalization Constant) = Number of correctly classified samples. For example, if a rule has *WtGain* count of 9, *WtLoss* count of 2, *RuleWeight* of 0.6, NC = 50, *UpdateConst* = 1, *NewRuleWeight* = 0.74. As is evident, the rule needed a net weight gain, and the new weight is greater than the old weight. Without NC, the new rule weight would have been 7.6, which is impractical, and will also greatly misclassify many currently correctly classified samples. The lower the value of the normalization constant, the higher would be the weight update. The more uncontrolled a weight update is, the more it will affect the already correctly classified samples. Hence, after much experimentation with many normalization constants for many datasets, it was found that the use of the total number of correctly classified samples as NC resulted in a better rule update process.

- 4.8 Store the final updated rules in *FinalRules*, which contains all the classification rules selected after six levels of pruning and rule updation. This associative classifier can be used for classifying new datasets.

5. WRUBAC evaluation: *FinalRules* to be used

- 5.1 **Evaluation using the test set:** For each sample in the test set, do the following:

- a. For each set of class rules: Add the *NewRuleWeight* of every rule whose LHS items are present in the current sample to get the cumulative rule weight for that class.
- b. Repeat step (a) for the remaining classes.
- c. Assign the class label that has the maximum cumulative weight as the predicted class for that sample.
- d. Repeat steps (a) to (c) for all the samples in the test set.
- e. Calculate and output the classification accuracy.

5.2 Evaluation using the whole dataset with cross-validation resampling technique:

- a. Split the dataset into ten folds using stratified sampling.
- b. Use one part as the data for each fold. For each sample in each fold, do the following:
 - i. For each set of class rules: Add the *NewRuleWeight* of every rule whose LHS items are present in the current sample to get the cumulative rule weight for that class.
 - ii. Repeat step (i) for the remaining classes.
 - iii. Assign the class label that has the maximum cumulative weight as the predicted class for that sample.
 - iv. Repeat steps (i) to (iii) for all the samples in the fold set.
 - v. Calculate the fold accuracy.
- c. Repeat step (b) for the next fold. Repeat until all ten folds have been evaluated.
- d. Report the final classification accuracy as the average of accuracies over all folds.

The above WRUBAC algorithm should be tested for the other selected feature subsets. For every dataset, there will be a combination of a feature subset, support percentage, confidence percentage, *InputThreshold*, and *UpdateConst* values that will result in the best possible associative classifier for that dataset. The *InputThreshold* parameter is mainly used to control pruning, whereas the *UpdateConst* is used to control the accuracy. The *InputThreshold* parameter should be used with caution. Only the value, which does not decrease the best acquired accuracy for the dataset using the

UpdateConst value, should be selected as the optimal *InputThreshold* value. A flowchart depicting an overview of the steps involved in WRUBAC is presented in Figure 6.1.

A note on the resampling technique used: Since a huge number of association rules are extracted, subjecting the dataset to ten-fold cross validation (for the purpose of rule extraction) would result in an additional rule merging step to consolidate the set of rules generated in every fold. This is a time consuming and computationally expensive task, especially in datasets with large number of samples or features and therefore, usually large number of rules. Hence, an apt way is to use the hold-out technique and split the dataset into 70% training set and 30% test set. In general, for the stratified hold-out resampling technique, the most commonly employed data proportion for the test set is one-third of the total number of available samples ($\approx 30\%$). The remaining two-third of the data is used for training (Soman et al., 2006). This is to ensure that there are sufficient amount of training samples for the classifier to extract rules. Since stratified technique is used, the proportion of the samples belonging to all the classes will be the same in the training and test sets, thereby ensuring that there is no class imbalance and hence, no biased final accuracy. The training set is used for rule mining and classifier development (rule pruning and weight update), and the test set is used for evaluating the built classifier. Additionally, to obtain more robust performance measures, the whole dataset (both test and training sets) was subjected to ten-fold cross validation based evaluation with the rules selected using the training set.

WRUBAC - development and evaluation

Contd. in next page

Figure 6.1 Flow chart depicting the WRUBAC algorithm

6.2 WRUBAC evaluation results

The WRUBAC algorithm was tested on eight datasets. The description of the benchmark datasets used and the WRUBAC evaluation results are presented in this section.

6.2.1 Description of the benchmark datasets

The benchmark datasets were downloaded from the UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/datasets.html>). The description of the datasets is presented in Table 6.1 for ease of comparison. In the case of the TTSH dataset, as highlighted in the previous chapter, the Post-BDS LOS alone was used as a feature for WRUBAC. In the case of the US dataset, the feature subsets *FFS1*, *FFS2*, *FFS3*, and *WFS*, were evaluated (Table 5.4, p.116).

Table 6.1 Description of the benchmark datasets

Name of the dataset	Hepatitis	Iris	Lung cancer	PIMA Indians diabetes	Wine	Breast Cancer Wisconsin (Original) (WOBC)
Original dataset	155 samples, 19 features + class, 2 classes	150 samples, 4 features + class, 3 classes	32 samples, 56 features + class, 3 classes	768 samples, 8 features + class, 2 classes	178 samples, 13 features + class, 3 classes	699 samples, 9 features + class, 2 classes
Feature: discretization intervals used (discretization done using WEKA “supervised discretize filter”)	1. Age: 10, 20, 30, 40, 50, 60, 70, 80			1. No. of times pregnant: 6.5	1. Alcohol: 12.78	1. Clump thickness
	2. Sex: Male, Female			2. Plasma glucose conc.: 99.5, 127.5, 154.5	2. Malic acid: 1.55, 2.455	2. Uniformity of cell size
	3. Steroid: No, Yes			3. Diastolic blood pressure: No reliable intervals	3. Ash: 2.03	3. Uniformity of cell shape
	4. Antiviral: No, Yes	1. Sepal length: 5.55, 6.15	No description provided for the 56 features;	4. Skin fold thickness: No reliable intervals	4. Alkalinity of ash: 17.9	4. Marginal adhesion
	5. Fatigue: No, Yes	2. Sepal width: 2.95, 3.35	Referred to by numbers 1-56;	5. Serum insulin: 14.5, 126.5	5. Magnesium: 88.5	5. Single epithelial cell size
	6. Malaise: No, Yes	3. Petal length: 2.45, 4.85	All predictive features are nominal, taking on integer values (0-3), so no discretization needed	6. BMI: 28.85	6. Total phenols: 1.84, 2.335	6. Bare nuclei
	7. Anorexia: No, Yes	4. Petal width: 0.8, 1.75		7. Diabetes pedigree: No reliable intervals	7. Flavonoids: 0.99, 1.575, 2.31	7. Bland chromatin
	8. Liver big: No, Yes	Class: Setosa, Versicolour, Virginica		8. Age: 28.5	8. Nonflavanoids: 0.425	8. Normal nucleoli
	9. Liver firm: No, Yes			Class: Negative, Positive For Diabetes	9. Proanthocyanins: 1.27	9. Mitoses
	10. Spleen palpable: No, Yes				10. Color intensity: 3.46, 7.55	10. Class (benign, malignant);
	11. Spiders: No, Yes				11. Hue: 0.855	All predictive features are nominal, taking on integer values (1-10), so no discretization needed
	12. Ascites: No, Yes				12. OD280: 2.115, 2.475	
	13. Varices: No, Yes				13. Proline: 476, 755, 987.5	
	14. Bilirubin: 0.39, 0.80, 1.20, 2.00, 3.00, 4.00				Class: three types of wines (1, 2, 3)	
	15. Sgot: 13, 100, 200, 300, 400, 500,					
	16. Albumin: 2.1, 3.0, 3.8, 4.5, 5.0, 6.0					
	17. Histology: No, Yes					
	18. Alk phosphate: 33, 80, 120, 160, 200, 250					
	19. Protime: 10, 20, 30, 40, 50, 60, 70, 80, 90					
	Class: Die, Live					

WRUBAC - development and evaluation

Name of the dataset	Hepatitis	Iris	Lung cancer	PIMA Indians diabetes	Wine	Breast Cancer Wisconsin (Original) (WOBC)
Missing values	Features ALK PHOSPHATE and PROTIME, and samples 42, 73, 84, 148, 57 were removed due to many missing values.	None	Missing values were replaced by the respective class averages	None	None	Some feature values missing in 16 samples. These samples were removed
Pre-processed dataset	150 samples, 17 features + class, 2 classes	150 samples, 4 features + class, 3 classes	32 samples, 56 features + class, 2 classes	768 samples, 8 features + class, 2 classes	178 samples, 13 features + class, 3 classes	683 samples, 9 features + class, 2 classes
Class distribution	Class <i>Die</i> : 31 Class <i>Live</i> : 119	Each class has 50 samples.	Class 1: 9 Class 2: 13 Class 3: 10	Class <i>Negative</i> : 500 Class <i>Positive</i> : 268	Class 1: 59 Class 2: 71 Class 3: 48	Class <i>Benign</i> : 444 Class <i>Malignant</i> : 239
Best feature subset selected after evaluating in WRUBAC	{12, 14, 16, 11}	{3,4}	{19, 20, 6, 24, 37}	{2,5,6,8}	{1,7}	{1,2,6}

6.2.2 Results and discussion

The evaluation results are presented in Table 6.2. The *UpdateConst* and *InputThreshold* parameters that were used to get the best results for each dataset are listed in the table. The resampling technique described under section 6.1 was employed. Consequently, five accuracy values are listed in the table: accuracy obtained using the non-updated rules on the 70% training dataset (A1), accuracy obtained using the updated rules on the 70% training dataset (A2), accuracy obtained using the updated rules on the 30% test dataset (A3), the overall accuracy obtained using the non-updated rules on the whole dataset (training and test samples put together) (A4), and the overall accuracy obtained using the updated rules on the whole dataset (A5).

For the Hepatitis, Iris, PIMA, Wine, and WOBC datasets, the obtained WRUBAC accuracy is compared with the accuracy presented by a CBA-like associative classifier that is based on the database coverage (DBC) pruning technique (Chiusano et al., 2009). Chiusano et al. evaluated these five datasets on the CBA-like classifier using the following pruning techniques: redundancy removal, confidence based pruning, database coverage, chi-squared test, and PER method. For all these five datasets, the database coverage technique resulted in the best accuracy. Hence, those results were used as comparison metrics in this work. In the case of the Lung dataset, the comparison is made with the results obtained by Aeberhard et al. (1994) using the regularized discriminant analysis classifier. This dataset was chosen because of its high dimensionality - 56 features and only 32 samples.

The results that are highlighted in gray indicate the datasets for which WRUBAC performed better than the literature method. The Hepatitis dataset, without feature selection, would have resulted in 55,436 rules when a minimum support threshold of 1% is used. The database coverage technique worked through these rules, selected 33 rules, and presented an accuracy of 86.45%. WRUBAC, on the other hand, worked with an initial rule set of just 384 rules, selected 41 rules, and presented an accuracy of 87.23% in just 8.72 seconds. The time and space savings are quite evident.

WRUBAC - development and evaluation

Table 6.2 WRUBAC algorithm evaluation results

Dataset	No. of samples/ No. of features (minus class attribute)	InputT hreshol d/ Update Const	A1*	A2*	A3*	A4*	A5*	AUC _4	AUC _5	Total number of rules after pruning level					Time (s)	The best results in the literature for the same dataset
										1	2,3	4	5	6		
Hepatitis	150/17	0.1/1	85.85	87.74	86.36	85.50	87.23	0.81	0.82	384	272	79	48	41	8.72	86.45% (DBC, Apriori, 33 rules) (Chiusano et al., 2009)
IRIS	150/4	0.1/1	95.24	97.14	93.33	95.33	96.00	NA	NA	27	27	9	9	9	2.75	93.34% (DBC, Apriori, 10 rules) (Chiusano et al., 2009)
Lung	32/56	0.1/1	87.50	87.50	75.00	82.50	85.00	NA	NA	2755	2340	414	195	191	14.78	62.5% (RDA algorithm - Aeberhard et al., 1994)
PIMA	768/8	0.4/1	78.07	78.81	77.83	77.99	78.52	0.78	0.78	505	296	99	53	53	48.21	78.52% (DBC, Apriori, 68 rules) (Chiusano et al., 2009)
TTSH dataset	149/9	0.2/1	93.80	93.80	96.23	94.51	94.51	0.97	0.97	8	6	4	4	4	2.51	94.51% (Original clinical trial results)
US dataset: FFS1/FFS2	241/34	0.4/1	92.94	94.12	76.06	88.01	88.83	0.82	0.83	554	232	141	109	109	25.56	Around 63% (Original clinical trial results)
US dataset: FFS3	241/34	0.4/3	92.94	95.29	80.28	89.22	90.86	0.81	0.87	1195	691	244	188	188	43.64	Around 63% (Original clinical trial results)
US dataset: WFS	241/34	0.4/7	88.24	90.59	71.83	83.46	85.00	0.74	0.80	431	192	118	72	72	18.85	Around 63% (Original clinical trial results)
Wine	178/13	0.2/1	94.44	94.44	96.15	94.90	95.00	NA	NA	57	44	18	12	12	3.59	98% (DBC, Apriori, 18 rules) (Chiusano et al., 2009)
WOBC	683/9	0.2/1	94.36	94.36	94.12	94.31	94.31	0.93	0.93	137	87	36	34	34	17.85	95.71% (DBC, Apriori, 63 rules) (Chiusano et al., 2009)

* where A1 = Accuracy obtained using the training dataset and non-updated rules (in %); A2 = Accuracy obtained using the training dataset and updated rules (in %)
A3 = Accuracy obtained using the test dataset and updated rules (in %); A4 = Accuracy obtained using the overall dataset and non-updated rules (cross- validation) (in %); A5 = Accuracy obtained using the overall dataset and updated rules (cross- validation) (in %); AUC_4 = AUC obtained using the overall dataset and non-updated rules; AUC_5 = AUC obtained using the overall dataset and updated rules; DBC = Database Coverage

A good accuracy improvement is seen in IRIS and Lung datasets. Even though the accuracy seems to be slightly lower for the Wine and WOBC datasets, a good amount of reduction in the number of rules can be observed. The PIMA dataset presented a similar accuracy with a lesser number of rules compared to the database coverage method.

By controlling the *InputThreshold* value, some more rules were pruned in the Hepatitis and Lung datasets, without compromising the accuracy. Infact the accuracy was improved (see respective A4 and A5 values). In Table 6.2, this pruning can be observed as a decrease in the number of rules between levels 5 and 6 of pruning. By changing the *UpdateConst* value, the best accuracy for a given dataset was determined. For example, in the case of the US dataset (*FFS3*), an AUC value of 0.84 was obtained for an *UpdateConst* value of 2 (not shown in Table 6.2), whereas a value of 0.87 was obtained for an *UpdateConst* value of 3. This establishes the utility of having the *UpdateConst* parameter for controlling the amount of rule updation.

Also, from Table 6.2, it can be observed that WRUBAC works well irrespective of the number of samples and the number of features. Even for a complex dataset like Lung dataset, which has 32 samples and 56 features, the algorithm presented good results.

To understand the effect of the rule updation system, the accuracy and AUC values obtained using the non-updated rules on a dataset (A4 and AUC_4) should be compared with those obtained using the updated rules (A5 and AUC_5) on the same dataset. It can be seen that A5 and AUC_5 are higher or equal to A4 and AUC_4 in all the cases. Another important observation is that the time taken to extract rules, apply the six pruning methods, evaluate both the training and test data, and also to evaluate the whole dataset with ten-fold cross validation is less than a minute for each datasets. Thus, it is evident that the WRUBAC algorithm can provide accuracies that are equal to and sometimes higher than the accuracies presented by the current associative classifiers with less time and memory requirements and lesser rules.

To illustrate the change in the weights after rule updation, the non-updated rules and updated rules obtained using the IRIS dataset are listed in Table 6.3. The rule updation process is carried out after pruning level 6. The IRIS dataset has nine rules before and

WRUBAC - development and evaluation

after pruning level 6, i.e. none of the rules were removed based on the RGT value (Table 6.3). Rule (33 => 3) with a weight of 0.9143 means *the third feature with a discretized code of 3* has a rule weight of 0.9143 for class 3. Rule (33, 43 => 3) with a weight of 0.8571 means the combination of *the third feature with a discretized code of 3* and *the fourth feature with a code of 3* has a rule weight of 0.8571 for class 3. Therefore, if the values 33 and 43 are both present in a sample (row), the sample is assigned a rule weight of 0.8571. Since the values 33 and 43 are also independently present in the sample, weights 0.9143 and 0.9143 are assigned. Overall, the cumulative weight is 2.2657 for class 3. If the values 31, 41, 32, and 42 are not present in this sample, then the cumulative weights for classes 1 and 2 would be zero. Therefore, based on class cumulative weights, this sample is categorized to belong to class 3.

Table 6.3 IRIS dataset: non-updated and updated rules

Non-updated rules (Pruned3Rules)						Updated rules (FinalRules)		
	LHS	RHS	RuleWeight	WtLoss	WtGain	LHS	RHS	NewRuleWeight
Class 1 rules	31	1	1	0	0	31	1	1
	41	1	1	0	0	41	1	1
	31, 41	1	1	0	0	31, 41	1	1
Class 2 rules	32	2	0.9143	3	0	32	2	0.8843
	42	2	1	3	0	42	2	0.97
	32, 42	2	0.9143	1	0	32, 42	2	0.9043
Class 3 rules	33	3	0.9143	0	2	33	3	0.9343
	43	3	0.9143	0	2	43	3	0.9343
	33, 43	3	0.8571	0	0	33, 43	3	0.8571

In Table 6.3, *WtLoss* and *WtGain* parameters indicate the number of times the rule has to lose or gain weight. These parameters are determined based on the misclassifications done by the rule (as explained in the WRUBAC algorithm section). For the IRIS dataset, class 1 rules need no further weight updation as they have correctly classified all the samples (*WtLoss* and *WtGain* = 0). Some class 2 and class 3 rules need updation as shown in the table. For example, the rule (32 => 2) needs a weight loss and therefore, the rule updation process has reduced its weight from 0.9143 to 0.8843. Rule (22 => 3), on the other hand, needs a weight gain, and hence, WRUBAC has improved its weight from 0.9143 to 0.9343.

WRUBAC - development and evaluation

WRUBAC did not improve the accuracy of the TTSH dataset and neither did the other classifiers used on this dataset. In the case of the US dataset, WRUBAC resulted in an accuracy of 88.83% for *FFS1*, and 85% for *WFS*. Based on the AUC values in Table 5.6 (p.118), BPNN and SVM classifiers rank high. For the ease of comparison of the results of WRUBAC with that of other classifiers (BPNN, SVM, and LDA/*WFS*), the respective performance measures are presented in Table 6.4.

Table 6.4 Comparison of WRUBAC, BPNN, SVM, and LDA/*WFS* classifier results for the US dataset

	FFS1/ FFS2			FFS3			WFS	
	BPNN	SVM	WRUBAC	BPNN	SVM	WRUBAC	LDA/ <i>WFS</i>	WRUBAC
Sensitivity (%)	74.14	63.79	72.33	75.86	67.24	78.00	93.10	70.33
Specificity (%)	91.80	97.27	93.98	92.90	96.17	95.03	98.36	89.62
Accuracy (%)	87.55	89.21	88.83	88.80	89.21	90.86	97.10	85.00
AUC	0.83	0.80	0.83	0.84	0.82	0.87	0.96	0.80
No. of rules	NA	NA	109	NA	NA	188	NA	72

From Table 6.4, it can be observed that WRUBAC resulted in an AUC similar to that of BPNN for the *FFS1/FFS2* subset. Moreover, it resulted in significantly higher AUC value (0.87) for the *FFS3* subset. However, in the case of the *WFS* subset, the AUC was only 0.80 as opposed to 0.97 given by the LDA/*WFS* classifier. It is a known fact that classifiers are highly data dependent, and some classifiers work better for some datasets. This may be the reason why LDA performed better here. Nevertheless, WRUBAC provided improved performance measures for the US dataset, even for the feature subset *WFS* that does not have the subjective Post-BDS LOS parameter.

6.3 Key features of the WRUBAC algorithm

In summary, the following are the key features of this algorithm:

1. Owing to the initial feature selection process, only the best features are chosen to form a reduced dataset. This greatly removes the redundant correlated features. Therefore, the dataset size is reduced making the rule extraction process fast. One may argue that the key feature of association rule mining, namely, the extraction of a very comprehensive rule base, is suppressed by this process. However, in practical terms, considering complex and large databases, this step would prove to

be effective. Also, from the results and by comparison with standard methods, it is evident that that this step did not affect the accuracy of prediction.

2. There are six levels of rule pruning employed in the algorithm. Therefore, the number of rules used to build the final classifier is minimal. Moreover, it is evident from the results that WRUBAC is even capable of achieving a better accuracy with lesser number of rules in some benchmark datasets. The lesser the number of rules, the better is the interpretability of the results.
3. Two of the pruning techniques used, namely, the *conflicting rule weight based pruning*, and the *rule goodness threshold based pruning*, are novel approaches used in WRUBAC.
4. By being parametric, because of the use of user-defined parameters like *InputThreshold*, and *UpdateConst*, the classifier can be customized to the studied dataset.
5. Most ARM classifiers do not handle rules with multiple class labels, also called as conflicting rules. They just prune these rules. However, in WRUBAC, these rules are studied, and only those rules with better discriminatory power than the others are kept.
6. The Rule Updation System (RUS) used effectively in the algorithm tunes the weight of the selected rules. It is a novel approach in the field of associative classifier development. The RUS updates the rule weights based on the classification results and also uses a normalization parameter to update the rules in such a way that there is no over-updation. Over-updation will result in misclassifying the already correctly classified samples.
7. Moreover, some associative classifiers use only a single rule to classify a sample, which may lead to overfitting. However, WRUBAC considers all the rules that are applicable to a sample, calculates cumulative class weights, and assigns the class that has the maximum cumulative weight.

CHAPTER 7: CONCLUSIONS AND FUTURE DIRECTIONS

7.1 Conclusions

Handling the complexity of cancer diagnosis, treatment, and prognosis would not have been possible without the availability of the currently improving imaging systems. Today's latest imaging systems play a major role in the primary cancer screening, in the diagnosis and characterization of lesions, staging and restaging, treatment selection and treatment progress monitoring, and in the determination of the cancer recurrence. In this research work, a comprehensive review of the presently available breast imaging modalities was done in order to understand the limitations of these devices, and thereby, to lay out the possible areas that warrant for further research and development. It is evident from this review that there is no single modality that can be used to detect breast cancer, with significant accuracy, in women of all ages, breast densities, and with different breast conditions such as cancerous, benign, and scar tissues, microcalcifications, metastatic foci etc. Other limiting factors are cost, test time, radiation exposure, image resolution, test environment conditions, and subjectivity of interpretation of results. It is highly impossible that a single modality be useful for all conditions. Hence, hybrid modalities (multi-modalities approaches) like PET-CT and SPECT-CT are being developed. These modalities have presented improved diagnostic capabilities for particular sub groups of "difficult to be imaged" patients - e.g. young women with dense breasts, women with non-palpable or small lesions and the like.

Moreover, adjunct modalities, which provide added confidence to the diagnosis made by primary modalities such as mammography, are finding their way into the market these days. One such adjunct device is the Biofield Diagnostic System (BDS). BDS is based on using skin surface electropotentials for breast cancer detection. Since there are a few other devices that use the tissue electrical characteristics for breast cancer detection, like EIT and EIS, a perspective review of the principle used and the devices exploiting these technologies was done in this work.

Conclusions & Future Directions

One of the key objectives of this work was to evaluate the efficiency of BDS in a clinical setting. For this purpose, a clinical study was designed, patients were recruited and evaluated using the BDS test in TTSH, and the results were analyzed. BDS recorded a high sensitivity of 96.23%, specificity of 93.80%, accuracy of 94.51%, and AUC of 0.972. However, it was observed that there was a slight subjectivity associated with the BDS interpretation methodology. To alleviate this problem, it was presumed that classifiers could be used to develop a better objective interpretation model, with the help of the recorded skin surface electropotential values as features. Since these features were not available in the TTSH dataset, the US dataset was evaluated.

The US BDS test recorded a sensitivity of 89.66%, low specificity of 54.64%, low accuracy of 63.07%, and AUC of 0.77, probably because of the older version of BDS device used for data collection. Again, the use of classifiers was proposed to improve these low performance measures. Consequently, a generalized data mining framework, which is an important contribution in this work, was presented and used on the two BDS datasets. Various feature subsets and classifiers were evaluated in the process. The results show that the classification accuracy of the US dataset is significantly better with the use of classifiers. Moreover, the selected features improved the BDS prediction methodology, making it possible to develop BDS into a more objective adjunct diagnostic modality. Although the results showed that BDS cannot replace mammography and ultrasound, BDS does not have to remain a dependent adjunct diagnostic modality, but can be used as an independent stand-alone adjunct diagnostic modality for breast cancer detection. In addition, the BDS test is non-invasive, radiation-free, cost-effective, and does not cause discomfort to the patient. It can be repeated as often as needed. The test results obtained are objective and acquired within a maximum of 15 minutes. Given the current performance of mammography and the high sensitivity and specificity of BDS, apart from its ability to give a good second opinion on the need for biopsies, BDS can be adopted as a secondary diagnostic tool in the standard examination pathway, when an abnormality is discovered.

Some of the studied classifiers, such as neural network based BPNN and PNN, and SVM are not transparent to the end user. FIS resulted in rules and was deemed more transparent than BPNN, PNN, and SVM. However, the accuracy provided by FIS was not relatively high compared to that recorded by the other classifiers. The LDA/WFS

classifier, which presented significant accuracy improvement, and whose feature subset showed promise in improving the objectivity of the BDS interpretation methodology, might tend to give different results if a dataset with equal number of benign and malignant samples was used. This difference may be attributed to the loss of linear separability. Therefore, in the course of search for a more transparent classifier that could work well for all datasets without the strict dependence on linear separability condition, a new classification algorithm named WRUBAC that is based on association rules was developed in this work. The WRUBAC algorithm is novel in many ways (as highlighted in the previous chapter), and is the key contribution of this research.

The WRUBAC algorithm was evaluated using six benchmark datasets and the two BDS datasets. It was observed that the algorithm could provide accuracies that are equal to and sometimes higher than the accuracies presented by the current associative classifiers with less time and memory requirements and lesser rules. Even though it did not present an accuracy as high as the LDA/WFS classifier for the US dataset, WRUBAC provided better performance measures for the US dataset, even for the feature subset *WFS* that does not have the subjective Post-BDS LOS parameter, when compared to the original 63.07% accuracy.

Because of the use of rules for classification, the WRUBAC algorithm has a better interpretability. Hence, the algorithm has good scope in the field of medical diagnostics where doctors prefer the information, which is used by a diagnostic modality to predict the presence or absence of a disease, to be transparent and comprehensible. Using WRUBAC, the domain knowledge can be extracted in the form of rules, and the rules can be used to develop a graphical user interface based classification tool that could improve the prediction accuracy of diagnostic devices.

7.2 Future directions

This work has established that BDS does have the potential to be a reliable adjunct modality, especially in confidently ruling out biopsies or for asking the patient to go for a short term follow-up. However, its capability can be further improved if the following limitations are addressed.

Conclusions & Future Directions

- Currently, the system is not intended to be used for screening purposes. It is used for the differential diagnosis of localized suspicious lesions. The field carcinogenesis concept can be exploited to develop BDS for screening purposes when women do not have visible symptoms. Even though there have been pilot studies (Crowe et al., 1995), which used non-directed electrode array (electrodes are placed in a circular fashion around the breast since there are no prior visible lumps) to measure the electropotentials independent of lesion location, more studies are required to establish BDS as a screening modality.
- The accuracy of the electrode placement is dependent on the operator as there are no BDS images to guide the operator for accurate electrode placement. The current software indicates misplaced electrodes after the data has been collected. However, it would be better if the software indicates any misplaced electrodes or electrodes with less skin contact at an earlier stage.
- Improved electrode design, the use of more electrodes to improve spatial resolution, and positioning of electrodes based on enhanced understanding of the electrophysiological characterization of benign and malignant breast tissues and surrounding breast epithelia may expand the diagnostic potential of this modality, especially for non-palpable lesions.
- More clinical studies should be conducted. The measured electropotential values for different breast conditions should be collected and studied independently. This would prove valuable for understanding the reason for false positives.
- With the use of the electropotential data collected from more clinical studies, the classifiers developed in this work can be subjected to more evaluations, and then be integrated into the current BDS device for better objective disease detection.
- As highlighted earlier, BDS has a capability to measure the net electrical activity related to cellular proliferation as demonstrated by its high positive correlation with Thymidine labeling index and Ki-67 antigen expression. This finding can be utilized to evaluate the tissue proliferation before, after and during cycles of cancer therapy (Faupel et al., 1997).

In this work, the efficiency of WRUBAC was compared with one associative classifier that used different pruning techniques. Evaluations on a wide range of datasets and comparison with other rule based classifiers are necessary.

REFERENCES

- Aeberhard, S. Coomans, D., De Vel, O. (1994). Comparative analysis of statistical pattern recognition methods in high dimensional settings. *Pattern Recognition*, 27(8), 1065-1077.
- Agrawal, R., Amielinski, T., Swami, A. (1993). Mining association rule between sets of items in large databases. In P. Buneman, S. Jajodia (Eds.), *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Washington, DC, 207-216.
- Agrawal, R., Srikant, R. (1994). Fast algorithms for mining association rule. In *Proceedings of the 20th International Conference on Very Large Data Bases*, Santiago, Chile, 487-499.
- Aha, D.W. (1998). Feature weighting for lazy learning algorithm. In H. Liu, H. Motoda (Eds.), *Feature extraction, construction and selection: a data mining perspective*. (pp. 13-32). Kluwer Academic Publishers.
- Amalu, W.C., Hobbins, W.B., Head, J.F., et al. (2006). Infrared imaging of the breast - an overview. In J.D.Bronzino (Ed.), *Biomedical Engineering Handbook*. (pp. 25.1-25.21). CRC Press.
- Antonie, M., Zaïane, O., Coman, A. (2003). Associative classifiers for medical images. *Mining Multimedia and Complex Data (Lecture Notes in Artificial Intelligence, 2797)*, 68-83.
- Antonie, M., Zaïane, O. (2004). An associative classifier based on positive and negative rules. In *Proceedings of the 9th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, Paris, France, 64-69.
- Archer, F., Gros, C. (1971). Classification thermographique des cancers mammaires. *Bulletin du Cancer*, 58, 351-362.
- Assenheimer, M., Laver-Moskovitz, O., Malonek, D., et al. (2001). The T-scan technology: electrical impedance as a diagnostic tool for breast cancer detection. *Physiological Measurement*, 22(1), 1-8.
- Athanasiou, A., Tardivon, A., Ollivier, L., et al. (2009). How to optimize breast ultrasound. *European Journal of Radiology*, 69(1), 6-13.
- Avril, N., Mather, S.J., Roylance, R. (2007). FDG-PET and PET/CT in Breast Cancer Staging. *Breast Care*, 2(6), 372-377.
- Baralis, E., Torino, P. (2002). A lazy approach to pruning classification rules. In *Proceedings of the 2002 IEEE International Conference on Data Mining*, Maebashi City, Japan, 35.

References

-
- Baralis, E., Chiusano, S., Graza, P. (2004). On support thresholds in associative classification. In *Proceedings of the 2004 ACM Symposium on Applied Computing*, Nicosia, Cyprus, 553-558.
- Baralis, E., Chiusano, S., Garza, P. (2008). A lazy approach to associative classification. *IEEE transactions on Knowledge and Data Engineering*, 20(2), 156-171.
- Bayford, R.H. (2006). Bioimpedance Tomography (Electrical Impedance Tomography). *Annual Review of Biomedical Engineering*, 8, 63-91.
- Berg, W.A., Blume, J.D., Cormack, J.B., et al. (2008). Combined screening with ultrasound and mammography vs. mammography alone in women at elevated risk of breast cancer. *Journal of American Medical Association*, 299(18), 2151-2163.
- BDS physician manual (LBL024 rev A and BDS Dossier463_617_389 rev 1).
- Blad, B., Baldetorp, B. (1996). Impedance spectra of tumor tissue in comparison with normal tissue: a possible clinical application for electrical impedance tomography. *Physiological Measurement*, 17, A105-A115.
- Blad, B., Wendel, P., Jonsson, M., et al. (1999). An electrical impedance index to distinguish between normal and cancerous tissues. *Journal of Medical Engineering and Technology*, 23(2), 57-62.
- Bluemke, D.A., Gatsonis, C.A., Chen, M.H., et al. (2004). Magnetic resonance imaging of the breast prior to biopsy. *Journal of American Medical Association*, 292(22), 2735-2742.
- Boone, J.M., Lindfors, K.K. (2006). Breast CT: potential for breast cancer screening and diagnosis. *Future Oncology*, 2(3), 351-356.
- Bowen, S.L., Wu, Y., Chaudhari, A.J., et al. (2009). Initial characterization of a dedicated breast PET/CT scanner during human imaging. *Nuclear Medicine*, 50(9), 1401-1408.
- Brenner, R.J., Parisky, Y. (2007). Alternative Breast-Imaging Approaches. *Radiologic Clinics of North America*, 45(5), 907-923.
- Burns, R.P. (1997). Image guided breast biopsy. *American Journal of Surgery*, 173(1), 9-11.
- Campbell, A.M., Land, D.V. (1992). Dielectric properties of female human breast tissue measured *in vitro* at 3.2 GHz. *Physics in Medicine and Biology*, 37(1), 193-210.
- Carney, P.A., Miglioretti, D.L., Yankaskas, B.C., et al. (2003). Individual and combined effects of age, breast density, and Hormone Replacement Therapy

References

- use on the accuracy of screening mammography. *Annals of Internal Medicine*, 138(3), 168-175.
- Carsten Riis, C., Lernevall, A., Sorensen, F.B., et al. (2005). 3D ultrasound-based evaluation of lesions in the uncompressed breast. In E. Ueno, T. Shiina, M. Kubota, K. Sawai (Eds.), *Research and Development in Breast Ultrasound*. (pp. 151-155). Tokyo: Springer.
- Castelo, R., Feelders, A.J., Siebes, A. (2001). MAMBO: Discovering Association Rules Based on Conditional Independencies. In *Proceedings of the 4th International Conference on Advances in Intelligent Data Analysis (Lecture Notes In Computer Science, 2189)*, 289-298.
- Chaudhary, S.S., Mishra, R.K., Swarup, A., et al. (1984). Dielectric properties of normal & malignant human breast tissues at radiowave & microwave frequencies. *Indian Journal of Biochemistry & Biophysics*, 21(1), 76-79.
- Chauveau, N., Hamzaoui, L., Rochaix, P., et al. (1999). Ex vivo discrimination between normal and pathological tissues in human breast surgical biopsies using bioimpedance spectroscopy. *Annals of New York Academy of Sciences*, 873, 42-50.
- Cheng, H.D., Cai, X., Chen, X., et al. (2003). Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern Recognition*, 36(12), 2967-2991.
- Cheng, H.D., Shi, X.J., Min, R., et al. (2006). Approaches for automated detection and classification of masses in mammograms. *Pattern Recognition*, 39(4), 646-668.
- Cheng, H.D., Shan, J., Ju, W., et al. (2010). Automated breast cancer detection and classification using ultrasound images: a survey. *Pattern Recognition*, 43(1), 299-317.
- Cherepenin, V., Karpov, A., Korjnevsky, et al. (2001). A 3D electrical impedance tomography (EIT) system for breast cancer detection. *Physiological Measurement*, 22, 9-18.
- Cherepenin, V.A., Karpov, A.Y., Korjnevsky, A.V., et al. (2002). Three-dimensional EIT imaging of breast tissues: system design and clinical testing. *IEEE Transactions in Medical Imaging*, 21(6), 662-667.
- Chiusano, S., Garza, P. (2009). Selection of high quality rules for associative classification. In Y. Zhao, C. Zhang, L. Cao (Eds.), *Post-mining of Association Rules: Techniques for Effective Knowledge Extraction*. (pp. 173-198). Information Science Reference.
- Crowe, S.S., Faupel, M.L. (1995). Use of non-directed (screening) arrays in the evaluation of symptomatic and asymptomatic breast patients. In J.M. Dixon (Ed.), *Electropotentials in the assessment of breast neoplasia*. (pp. 52-62). Berlin: Springer Verlag.

References

-
- Cuzick, J., Holland, R., Barth, V., et al. (1998). Electropotential measurements as a new diagnostic modality for breast cancer. *Lancet*, 352(9125), 359-363.
- Cuzick, J., Davies, R., Stephens, J., Borgwardt, V., Housworth, C., Lei, X., Gadsby, P., and Iskac, A. (2002). inventors; Biofield Corp., assignee. Method and apparatus for sensing and processing biopotentials. United States patent US 6351666.
- Davies, R.J. (1996). Underlying mechanism involved in surface electrical potential measurements for the diagnosis of breast cancer: an electrophysiological approach to breast cancer. In J.M. Dixon (Ed.), *Electropotentials in the clinical assessment of breast neoplasia*. (pp. 4-17). New York: Springer Verlag.
- Dickhaut, M., Schreer, I., Frischbier, H.J., et al. (1996). The value of BBE in the assessment of breast lesions, In J. M. Dixon (Ed.), *Electropotentials in the clinical assessment of breast neoplasia*. (pp. 63-69). New York: Springer Verlag.
- Diebold, T., Jacobi, V., Scholz, B., et al. (2005). Value of Electrical Impedance Scanning (EIS) in the evaluation of BI-RADS III/IV/V-lesions. *Technology in Cancer Research and Treatment*, 4(1), 93-97.
- Dirisamer, A., Halpern, B.S., Flöry, D., et al. (2009). Integrated contrast-enhanced diagnostic whole-body PET/CT as a first-line restaging modality in patients with suspected metastatic recurrence of breast cancer. *European Journal of Radiology*, doi:10.1016/j.ejrad.2008.10.031.
- Dua, S., Singh, H., Thompson, H.W. (2009). Associative classification of mammograms using weighted rules. *Expert Systems with Applications*, 36(5), 9250-9259.
- Elter, M., Horsch, A. (2009). CADx of mammographic masses and clustered microcalcifications: a review. *Medical Physics*, 36(6), 2052-2068.
- Faupel, M.L., Hsu, Y.-S. (1995). Dedicated systems for surface electropotential evaluation in the detection and diagnosis of neoplasia. In J. M. Dixon (Ed.), *Electropotentials in the clinical assessment of breast neoplasia*. (pp. 37-44). New York: Springer Verlag.
- Faupel, M.L., Vanel, D., Barth, V., et al. (1997). Electropotential evaluation as a new technique for diagnosing breast lesions. *European Journal of Radiology*, 24(1), 33-38.
- Faupel, M., Barrett, B., Stephens, J., Nathanson, S. inventors; (1998). Biofield Corp., assignee. D.C biopotential sensing electrode and electroconductive medium for use therein. United States patent US 5823957.
- Fayyad, U.M, Irani, K.B. (1993). Multi-Interval discretization of continuous-valued attributes for classification learning, In R. Bajcsy (Ed.), *Proceedings of the 13th*

References

-
- International Joint Conference on Artificial Intelligence*, Chambéry, France, 1022-1029.
- Fear, E.C., Hagness, S.C., Meaney, P.M. et al. (2002). Enhancing breast tumor detection with near-field imaging. *IEEE Microwave Magazine*, 3(1), 48-56.
- Foster, K.R., Schwan, H.P. (1989). Dielectric properties of tissues and biological materials: a critical review. *Critical Reviews in Biomedical Engineering*, 17(1), 25-104.
- Foster, J.J., Barkus, E., Yavorsky, C. (2005). *Understanding and Using Advanced Statistics: A Practical Guide for Students*. Sage Publications Ltd.
- Fricke, H., Morse, S. (1926). The electric capacity of tumors of the breast. *Journal of Cancer Research*, 16, 310-376.
- Fuchsjaeger, M.H., Flory, D., Reiner, C.S., et al. (2005). The negative predictive value of electrical impedance scanning in BI-RADS category IV breast lesions. *Investigative Radiology*, 40(7), 478-485.
- Fukuda, M., Shimizu, K., Okamoto, N., et al. (1996). Prospective evaluation of skin surface electropotentials in Japanese patients with suspicious breast lesions. *Japanese Journal of Cancer Research*, 87(10), 1092-1096.
- Gallager, H. S., Martin, J. E. (1969). Early phases in the development of breast cancer. *Cancer*, 24(6), 1170-1178.
- Garra, B.S., (2007). Imaging and estimation of tissue elasticity by ultrasound. *Ultrasound Quarterly*, 23(4), 255-268.
- Gatzemeier, W., Cuzick, J., Scelsi, M., et al. (2002). Correlation between the Biofield diagnostic test and Ki-67 labeling index to identify highly proliferating palpable breast lesions. *Breast Cancer Research and Treatment*, 76, S114.
- Germer, U., Tetzlaff, A., Geipel, A., et al. (2002). Strong impact of estrogen environment on Doppler variables used for differentiation between benign and malignant breast lesions. *Ultrasound in Obstetrics and Gynecology*, 19(4), 380-385.
- Gibson, A., Dehghani, H. (2009). Diffuse optical imaging. *Philosophical Transactions of the Royal Society A - Mathematical, Physical and Engineering Sciences*, 367(1900), 3055-3072.
- Gilhuija, G.A.K. (2008). Breast lesions: Computerized analysis of magnetic resonance imaging. In M.A. Hayat (Ed.), *Lung and breast carcinomas*. (pp. 529-538). Elsevier.
- Glickman, Y.A., Filo, O., Nachaliel, U., et al. (2002). Novel EIS postprocessing algorithm for breast cancer diagnosis. *IEEE Transactions in Medical Imaging*, 21(6), 710-712.

References

-
- Goller, D.A., Weidema, W.F., Davies, R.J. (1986). Transmural electrical potential difference as an early marker in colon cancer. *Archives of Surgery*, 121(3), 345-350.
- Grosenick, D., Moesta, K.T., Möller, M., et al. (2005). Time-domain scanning optical mammography: I. Recording and assessment of mammograms of 154 patients. *Physics in Medicine and Biology*, 50(11), 2429-2449.
- Guo, Z., Siew, K.W., August, D.A., et al. (2006). Optical imaging of breast tumor through temporal log-slope difference mappings. *Computers in Biology and Medicine*, 36(2), 209-223.
- Hall, M., Frank, E., Holmes, G., et al. (2009). The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, 11(1).
- Han, J., Pei, J., Yin, Y. (2000). Mining frequent patterns without candidate generation. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, Dallas, Texas, 1-12.
- Han, J., Pei, J., Yin, Y., et al. (2004). Mining frequent patterns without candidate generation: A Frequent-Pattern Tree approach, *Data Mining and Knowledge Discovery*, 8(1), 53-87.
- Hashimoto, B.E. (2008). *Practical Digital Mammography*. Thieme Medical Publishers.
- Hendrick, E.R. (2007). *Breast MRI: Fundamentals and Technical Aspects*. Springer.
- Heywang-Köbrunner, S.H., Viehweg, P., et al. (1997). Contrast-enhanced MRI of the breast: accuracy, value, controversies, solutions. *European Journal of Radiology*, 24(2), 94-108.
- Holder, S. (2005). *Electrical Impedance Tomography-Methods, History and Applications*. Bristol and Philadelphia: Institute of Physics Publishing.
- Horner, M.J., Ries, L.A.G., Krapcho, M., et al. (2009). *SEER Cancer Statistics Review, 1975-2006*. Bethesda, MD: National Cancer Institute.
- Hospers, G.A., Helmond, F.A., de Vries, E.G., et al. (2008). PET imaging of steroid receptor expression in breast and prostate cancer. *Current Pharmaceutical Design*, 14(28), 3020-3032.
- Huang, W., Fisher, P.R., Dulaimy, K., et al. (2004). Detection of breast malignancy: diagnostic MR protocol for improved specificity. *Radiology*, 232(2), 585-591.
- Itoh, A., Ueno, E., Tohno, E., et al. (2006). Breast disease: clinical application of US elastography for diagnosis. *Radiology*, 239(2), 341-350.

References

-
- John, G.H., Kohavi, R., Pfleger, K. (1994). Irrelevant Features and the Subset Selection Problem. In *Proceedings of the 11th International Conference on Machine Learning*, San Mateo, California, Morgan Kaufmann, 121-129.
- Joines, W.T., Zhang, Y., Li, C., et al. (1994). The measured electrical properties of normal and malignant human tissues from 50 to 900 MHz. *Medical Physics*, 21(4), 547-550.
- Jossinet, J. (1996). Variability of impedivity in normal and pathological breast tissue. *Medical and Biological Engineering and Computing*, 34(5), 346-350.
- Jossinet, J. (1998). The impedivity of freshly excised human breast tissue. *Physiological Measurement*, 19(1), 61-75.
- Jossinet, J., Schmitt, M. (1999). A review of parameters for the bioelectrical characterization of breast tissue. *Annals of New York Academy of Sciences*, 873(1), 30-41.
- Joy, J.E., Penhoet, E.E., Petitti, D.B. (2005). *Saving women's lives - Strategies for improving breast cancer detection and diagnosis*. Washington, D.C: The National Academic Press.
- Katzenellenbogen, J.A., Welch, M.J., Dehdashti, F. (1997). The development of estrogen and progestin radiopharmaceuticals for imaging breast cancer. *Anticancer Research*, 17(3B), 1573-1576.
- Keith, L.G., Oleszczuk, J.J., Laguens, M. (2001). Circadian rhythm chaos: A new breast cancer marker. *International Journal of Fertility and Women's Medicine*, 46 (6), 238-247.
- Keyserlingk, J.R., Ahlgren, P.D., Yu, E., et al. (2008). Functional infrared imaging of the breast: Historical perspectives, current application and future considerations. In N.A. Diakides, J.D. Bronzino (Eds.), *Biomedical Engineering Handbook*. (pp. 10.1-10.30). CRC Press.
- Kira, K., Rendell, L. (1992). A practical approach to feature selection. In D. Sleeman, P. Edwards (Eds.), *International Conference on Machine Learning*, 368-377.
- Kook, S.H., Park, H.W., Lee, Y.R., et al. (1999). Evaluation of solid breast lesions with power Doppler sonography. *Journal of Clinical Ultrasound*, 27(5), 231-237.
- Korotkova, M., and Karpov, A. (2007). Procedure for assessment of the mammary gland electrical impedance images. In *Proceedings of the 13th International Conference on Electrical Bioimpedance and the 8th Conference on Electrical Impedance Tomography*, Graz, Austria, 699-702.
- Kuhl, C.K., Schrading, S., Leutner, C.C., et al. (2005). Mammography, Breast Ultrasound, and Magnetic Resonance Imaging for Surveillance of Women at

References

- High Familial Risk for Breast Cancer. *Journal of Clinical Oncology*, 23(33), 8469-8476.
- Kumar, R., Chauhan, A., Zhuang, H., et al. (2006). Clinicopathologic factors associated with false negative FDG-PET in primary breast cancer. *Breast Cancer Research and Treatment*, 98(3), 267-274.
- Lagalla, R., Midiri, M. (1998). Image quality control in breast ultrasound. *European Journal of Radiology*, 27(S2), S229-S233.
- Lavayssière, R., Cabée, A.E., Filmont, J.E. (2009). Positron Emission Tomography (PET) and breast cancer in clinical practice. *European Journal of Radiology*, 69(1), 50-58.
- Lazebnik, M., McCartney, L., Popovic, D., et al. (2007a). A large-scale study of the ultra wideband microwave dielectric properties of normal breast tissue obtained from reduction surgeries. *Physics in Medicine and Biology*, 52(10), 2637-2656.
- Lazebnik, M., Popovic, D., McCartney, L., et al. (2007b). A large-scale study of the ultrawideband microwave dielectric properties of normal, benign, and malignant breast tissues obtained from cancer surgeries. *Physics in Medicine and Biology*, 52(10), 6093-6115.
- Leff, D.R., Warren, O.J., Enfield, L.C., et al. (2008). Diffuse optical imaging of the healthy and diseased breast: A systematic review. *Breast Cancer Research and Treatment*, 108(1), 9-22.
- Lehman, C.D., Gatsonis, C., Kuhl, C.K., et al. (2007). MRI Evaluation of the Contralateral Breast in Women Recently Diagnosed with Breast Cancer. *The New England Journal of Medicine*, 356(13), 1295-1303.
- Lerman, H., Lievshitz, G., Zak, O., et al. (2007). Improved sentinel node identification by SPECT/CT in overweight patients with breast cancer. *Journal of Nuclear Medicine*, 48(2), 201-206.
- Li, W., Han, J., Pei, J. (2001). CMAR: Accurate and efficient classification based on multiple-class association rule. In *Proceedings of the International Conference on Data Mining*, San Jose, CA, 369-376.
- Lindfors, K.K., Boone, J.M., Nelson, T.R., et al. (2008). Dedicated breast CT: initial clinical experience. *Radiology*, 246(3), 725-733.
- Lister, D., Evans, A.J., Burrell, H., et al. (1998). The accuracy of breast ultrasound in the evaluation of clinically benign discrete, symptomatic breast lumps. *Clinical Radiology*, 53(7), 490-492.
- Liu, P.F., Debatin, J.F., Caduff, R.F., et al. (1998). Improved diagnostic accuracy in dynamic contrast enhanced MRI of the breast by combined quantitative and qualitative analysis. *The British Journal of Radiology*, 71(845), 501-509.

References

-
- Liu, H., Motoda, H. (1998). *Feature Selection for Knowledge Discovery and Data Mining*. MA: Kluwer Academic Publishers.
- Liu, B., Hsu, W., Ma, Y. (1998). Integrating classification and association rule mining. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, New York, 80-86.
- Liu, B., Ma, Y., Wong, C.-K. (2000). Improving an association rule based classifier. In *Proceedings of the 4th European Conference on Principles of Data Mining and Knowledge Discovery*, Lyon, France, 504-509.
- Lumachi, F., Zucchetta, P., Marzola, M.C. et al. (2002). Positive predictive value of 99mTc sestamibi scintimammography in patients with non-palpable, mammographically detected, suspicious, breast lesions. *Nuclear Medicine Communications*, 23(11), 1073-1078.
- Malich, A., Fritsch, T., Anderson, R., et al. (2000). Electrical impedance scanning for classifying suspicious lesions: first results. *European Radiology*, 10(10), 1555-1561.
- Malich, A., Bohm, T., Facius, M., et al. (2001a). Additional value of electrical impedance scanning: experience of 240 histologically-proven breast lesions. *European Journal of Cancer*, 37(18), 2324-2330.
- Malich, A., Boehm, T., Facius, M., et al. (2001b). Differentiation of mammographically suspicious lesions: evaluation of breast ultrasound, MRI mammography and electrical impedance scanning as adjunctive technologies in breast cancer detection. *Clinical Radiology*, 56(4), 278-283.
- Malich, A., Bohm, T., Facius, M., et al. (2003a). Electrical impedance scanning as a new imaging modality in breast cancer detection - a short review of clinical value on breast application, limitations and perspectives. *Nuclear Instruments and Methods in Physics Research A*, 497(1), 75-81.
- Malich, A., Facius, M. (2003b). Influence of size and depth on accuracy of electrical impedance scanning. *European Radiology*, 13(11), 2441-2446.
- Marino, A.A., Iliev, I.G., Schwalke, M.A., et al. (1994). Association between cell membrane potential and breast cancer. *Tumour Biology*, 15(2), 82-89.
- Martin, G., Martin, R., Brieva, M.J., et al. (2002). Electrical impedance scanning in breast cancer imaging: correlation with mammographic and histological diagnosis. *European Radiology*, 12(6), 1471-1478.
- Meisamy, S., Bolan, P.J., Garwood, M. (2008). Magnetic resonance spectroscopy of breast cancer: current techniques and clinical applications. In M.A. Hayat (Ed.), *Lung and breast carcinomas*. (pp. 407-415). Elsevier.
- Metz, C.E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 8, 283-298.

References

-
- Mills, F.C. (1965). *Statistical Methods*. London: Sir Isaac Pitnam and Sons.
- Morimoto, T., Kinouchi, Y., Iritani, T., et al. (1990). Measurement of the electrical bio-impedance of breast tumors. *European Surgical Research*, 22(2), 86-92.
- Morimoto, T., Kimura, S., Konishi, Y., et al. (1993). A study of the electrical bioimpedance of tumors. *Journal of Investigative Surgery*, 6(1), 25-32.
- Moses, W.W. (2001). Trends in PET imaging. *Nuclear Instruments and Methods in Physics Research*, A471, 209-214.
- Ng, E.Y. -K., Sudharsan, N.M. (2001). An improved 3-D direct numerical modeling and thermal analysis of a female breast with tumour. *International Journal of Engineering in Medicine*, 215(1), 25-37.
- Ng, E.Y. -K. (2009). A review of thermography as promising non-invasive detection modality for breast tumor. *International Journal of Thermal Sciences*, 48(5), 849-859.
- Nofech-Mozes, S., Mackenzie, R., Kahn, H.J., et al. (2008). Breast metastasis by medullary thyroid carcinoma detected by FDG positron emission tomography. *Annals of Diagnostic Pathology*, 12(1), 67-71.
- Orel, S.G., Schnall, M.D. (2001). MR Imaging of the Breast for the Detection, Diagnosis, and Staging of Breast Cancer. *Radiology*, 220(1), 13-30.
- Quaresima, V., Matcher, S.J., Ferrari, M. (1998). Identification and Quantification of Intrinsic Optical Contrast for Near-Infrared Mammography. *Photochemistry and Photobiology*, 67(1), 4 -14.
- Quinlan, J.R. (1993). *C4.5: Programs for machine learning*. San Fransisco, CA: Morgan Kaufmann Publishers Inc.
- Paliwal, M, Kumar, U.A. (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, 36(1), 2-17.
- Park, S.H., Goo, J.M., Jo, C.H. (2004). Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean Journal of Radiology*, 5(1), 11-18.
- Park, J.M., Ikeda, D.M. (2006). Promising techniques for breast cancer detection, diagnosis, and staging using non-ionizing radiation imaging techniques. *Physica Medica*, 21(S1), 7-10.
- Pasquier, N., Bastide, Y., Taouil, R. et al. (1999). Efficient mining of association rules using closed itemset lattices. *Information Systems*, 24 (1), 25-46.
- Peat, J, Barton, B. (2005). *Medical Statistics: A Guide to Data Analysis and Critical Appraisal*. BMJ Books.

References

-
- Pei, J., Han, J., Mao, R. (2000). CLOSET: An efficient algorithm for mining frequent closed itemsets. In *Proceedings of the 2000 ACM-SIGMOD International Workshop Data Mining and Knowledge Discovery*.
- Piperno, G., Frei, E.H., Moshitzky, M. (1990). Breast cancer screening by impedance measurements. *Frontiers of Medical and Biological Engineering*, 2(2), 111-117.
- Piperno, G., Lenington, S. (2002). Breast electrical impedance and estrogen use in postmenopausal women. *Maturitas*, 41(1), 17-22.
- Perlet, C., Kessler, M., Lenington, S., et al. (2000). Electrical impedance measurement of the breast: effect of hormonal changes associated with the menstrual cycle. *European Radiology*, 10(10), 1550-1554.
- Pethig, R. (1984). Dielectric properties of biological materials: Biophysical and medical applications. *IEEE Transactions on Electrical Insulation*, EI-19, 453-473.
- Pethig, R. (1987). Dielectric properties of body tissues. *Clinical Physics and Physiological Measurement*, 8A, 5-12.
- Pisano, E.D., Gatsonis, C., Hendrick, E., et al. (2005). Diagnostic performance of digital versus film mammography for breast-cancer screening. *New England Journal of Medicine*, 353(17), 1773-1783.
- Pisano, E.D., Hendrick, R.E., Yaffe, M.J., et al. (2008). Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology*, 246(2), 376-383.
- Polk, C., Postow, E. (1986). *CRC Handbook of Biological Effects of Electromagnetic Fields*. Boca Raton, Florida: CRC Press.
- Rigaud, B., Morucci, J.P., Chauveau, N. (1996). Bioelectrical impedance techniques in medicine. Part I: Bioimpedance measurement. Second section: impedance spectrometry. *Critical Reviews in Biomedical Engineering*, 24(4-6), 257-351.
- Sacchini, V (1996). Report of the European School of Oncology Task Force on Electropotentials in the clinical assessment of neoplasia. *The Breast*, 5(4), 282-286.
- Saslow, D., (2007). American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *A Cancer Journal for Clinicians*, 57(2), 75-89.
- Scaperrotta, G., Ferranti, C., Costa, C., et al. (2008). Role of sonoelastography in non-palpable breast lesions. *European Radiology*, 18(11), 2381-2389.
- Schillaci, O., Scopinaro, F., Danieli, R., et al. (1997). ⁹⁹Tcm-sestamibi scintimammography in patients with suspicious breast lesions: comparison of

References

-
- SPET and planar images in the detection of primary tumors and axillary lymph node involvement. *Nuclear Medicine Communications*, 18(9), 839-845.
- Schillaci, O., Buscombe, J.R. (2004). Breast scintigraphy today indications and limitations. *European Journal of Nuclear Medicine and Molecular Imaging*, 31(1), S35-45.
- Schilling, K., Narayanan, D., Kalinyak, J. E. (2008). Effect of Breast Density, Menopausal Status, and Hormone Use in High Resolution Positron Emission Mammography. *Radiological Society of North America*, VB31-04.
- Scholz, B., and Anderson, R. (2000). On electrical impedance scanning-principles and simulations. *Electromedica*, 68, 35-44.
- Schnall, M.D. (2003). Breast MR imaging. *Radiologic Clinics of North America*, 41(1), 43-50.
- Schulz-Wendtland, R., Fuchsjägerb, M., Wackerc, T., et al. (2009). Digital mammography: An update. *European Journal of Radiology*, 72(2), 258-265.
- Schwan, H.P. (1957). Electrical properties of tissue and cell suspensions. *Advances in Biomedical Engineering and Medical Physics*, 5, 147-209.
- Schwan, H.P., Takashima, S. (1993). Electrical conduction and dielectric behavior in biological systems. *Encyclopedia of Applied Physics*, 177-200.
- Scopinaro, F., Schillaci, O, Ussof, W., et al. (1997). A three center study on the diagnostic accuracy of 99mTc-MIBI scintimammography. *Anticancer Research*, 17(3B), 1631-1634.
- Seo, B.K., Pisano, E.D., Cho, K.R., et al. (2005). Low-dose multidetector dynamic CT in the breast: preliminary study. *Journal of Clinical Imaging*, 29(3), 172-178.
- Singapore cancer statistics. (2009). *Trends in Cancer Incidence in Singapore 2003-2007*. Singapore: National Cancer Centre.
- Singh, B., Smith, C.W., Hughes, R. (1979). *In vivo* dielectric spectrometer. *Medical and Biological Engineering and Computing*, 17(1), 45-60.
- Singh, V., Saunders, C., Wylie, L. (2008). New diagnostic techniques for breast cancer detection. *Future Oncology*, 4(4), 501-513.
- Sivaramakrishna, R., Powell, K.A., Lieber, M.L, et al. (2002). Texture analysis of lesions in breast ultrasound images. *Computerized Medical Imaging and Graphics*, 26(5), 303-307.
- Soman, K.P., Shyam, D., Ajay, V. (2006). *Insight into Data Mining: Theory and Practice*. India: Prentice Hall.
- Sotskova, N., Karpov, A., Korotkova, M., et al. (2007). Particularities of electrical impedance images in different forms of growth of infiltrative breast cancer. In

References

-
- Proceedings of the 13th International Conference on Electrical Bioimpedance and the 8th Conference on Electrical Impedance Tomography*, Graz, Austria, 560-563.
- Standards Technical Reference (STR 2003) for Thermal Imagers for Human Temperature Screening. Part 1: Requirements and Test Methods, TR 15-1, Spring Singapore, ISBN 9971-67-963-9.
- Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Stojadinovic A, Nissan, A., Gallimidi, Z., et al. (2005). Electrical impedance scanning for the early detection of breast cancer in young women: preliminary results of a multicenter prospective clinical trial. *Journal of Clinical Oncology*, 23(12), 2703-2715.
- Stojadinovic, A., Moskovitz, O., Gallimidi, Z., et al. (2006). Prospective study of electrical impedance scanning for identifying young women at risk for breast cancer. *Breast Cancer Research and Treatment*, 97(2), 179-189.
- Stojadinovic, A., Nissan, A., Shriver, C.D., et al. (2008). Electrical impedance scanning as a new breast cancer risk stratification tool for young women. *Journal of Surgical Oncology*, 97(2), 112-120.
- Surowiec, A.J., Stuchly, S.S., Barr, J.R., et al. (1988). Dielectric properties of breast carcinoma and the surrounding tissues. *IEEE Transactions on Biomedical Engineering*, 35(4), 257-263.
- Szabo, B.K., Saracco, A., Wilczek, B., et al. (2005). Adjunctive diagnostic value of targeted electrical impedance imaging to conventional methods in the evaluation of breast lesions. *Acta Radiology*, 46(8), 782-790.
- Tartar, M., Comstock, C.E., Kipper, M.S. (2008). *Breast cancer imaging: a multidisciplinary, multimodality approach*. Elsevier Health Sciences.
- Thabtah, F., Cowling, P., Peng, Y. (2004). MMAC: A new multi-class, multi-label associative classification approach. In *Proceedings of the 4th IEEE International Conference on Data Mining*, Brighton, UK, 217-224.
- Toivonen, H. Klemetinen, M., Ronkainen, P., et al. (1995). Pruning and grouping discovered association rules. In *MLnet Workshop on Statistics, Machine Learning, and Discovery in Databases*, 47-52.
- Van der Ploeg, I.M.C., Valdes Olmos, R.A., Nieweg, O.E., et al. (2007). The additional value of SPECT/CT in lymphatic mapping in breast cancer and melanoma. *Journal of Nuclear Medicine*, 48(11), 1756-1760.
- Van Houten, E.E.W., Doyley, M.M., Kennedy, F.E., et al. (2003). Initial *in vivo* experience with steady-state subzone-based MR elastography of the human breast. *Journal of Magnetic Resonance Imaging*, 17(1), 72-85.

References

-
- Vargas, H.I., Vargas, M.P., Gonzalez, K.D., et al. (2004). Outcomes of sonography-based management of breast cysts. *American Journal of Surgery*, 188(4), 443-447.
- Warner, E., Plewes, D.B., Shumak, R.S., et al. (2001). Comparison of Breast Magnetic Resonance Imaging, Mammography and Ultrasound for Surveillance of Women at High Risk for Hereditary Breast Cancer. *Journal of Clinical Oncology*, 19(15), 3524-3531.
- Wasserman, P.D. (1993). *Advanced Methods in Neural Computing*. New York: Van Nostrand Reinhold.
- Weinberg, I.N. (2006). Applications for positron emission mammography. *Physica Medica*, 21(1), 132-137.
- Weiss, B.A., Ganpola, G.A.P., Freeman, H.P., et al. (1994). Surface electrical potentials as a new modality in the diagnosis of breast lesions: a preliminary report. *Breast disease*, 7(2), 91-98.
- Wu, T., Stewart, A., Stanton, M., et al. (2003). Tomographic mammography using a limited number of low-dose cone-beam projection images. *Medical Physics*, 30(3), 365-380.
- Zaki, M.J., Parthasarathy, S., Ogihara, M. et al. (1997). New algorithms for fast discovery of association rules. In *3rd International Conference on Knowledge Discovery and Data Mining*, Newport Beach, California, 283-286.
- Zaki, M.J., Ching-Jui Hsiao. (2002). CHARM: An Efficient Algorithm for Closed Itemset Mining. In *2nd SIAM International Conference on Data Mining*, Arlington, VA, 457-473.
- Zaki, M.J., Gouda, K. (2003). Fast Vertical Mining Using Diffsets. In *9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, D.C., 326-335.
- Zakhireh, J., Gomez, R., Esserman, L. (2008). Converting evidence to practice: A guide for the clinical application of MRI for the screening and management of breast cancer. *European Journal of Radiology*, 44(18), 2742-2752.
- Zhenyu, J., Xiuzhen, D., Feng, F., et al. (2005). Study of influencing factors on Electrical Impedance Scanning imaging. In *Proceedings of the 27th Annual Conference of the IEEE Engineering in Medicine and Biology Society*, Shanghai, China, 2910-2913.
- Zweig, M.H., Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39(8), 561-577.

APPENDIX A: BREAST CANCER TYPES

Lobular carcinoma in situ (LCIS, lobular neoplasia): The term *in situ* refers to cancer that has not spread past the area where it initially developed. LCIS is a sharp increase in the number of cells within the lobules of the breast. Many physicians do not classify LCIS as breast cancer, and these patients are closely monitored every four months.

Ductal carcinoma in situ (DCIS): DCIS, the most common type of non-invasive breast cancer, is confined to the ducts of the breast. DCIS is often first detected on mammogram as microcalcifications. With early detection, the five-year survival rate for DCIS is nearly 100%, provided that the cancer has not spread past the milk ducts to the fatty breast tissue or any other regions of the body.

Adenocarcinoma: This is a general type of carcinoma that starts in glandular tissues anywhere in the body, and it accounts for nearly all breast cancers.

Infiltrating lobular carcinoma (ILC): ILC, also known as invasive lobular carcinoma, begins in the lobules of the breast, and often spreads to other regions of the body. ILC accounts for 10% to 15% of breast cancers.

Infiltrating ductal carcinoma (IDC): IDC, also known as invasive ductal carcinoma, begins in the milk ducts, penetrates the wall of the duct, and invades the fatty tissue of the breast and possibly other regions of the body. IDC is the most common type of breast cancer, accounting for about 80% of breast cancer diagnoses.

Medullary carcinoma: Medullary carcinoma is an invasive breast cancer that forms a distinct boundary between tumor tissue and normal tissue. Only 5% of breast cancers are medullary carcinoma.

Mucinous carcinoma: Also called *colloid carcinoma*, mucinous carcinoma is a rare breast cancer formed by the mucus-producing cancer cells. Women with mucinous carcinoma generally have a better prognosis than women with more common types of invasive carcinoma.

Tubular carcinoma: Tubular carcinomas are a special type of infiltrating breast carcinoma. Women with tubular carcinoma generally have a better prognosis than women with more common types of invasive carcinoma. Tubular carcinomas account for around 2% of breast cancer diagnoses.

Inflammatory breast cancer: Inflammatory breast cancer is characterized by the appearance of inflamed breasts with dimples and/or thick ridges caused by cancer cells blocking lymph vessels or channels in the skin over the breast. Though inflammatory breast cancer is rare (accounting for only 1% of breast cancers), it is extremely fast-growing.

Paget's disease of the nipple: This is a rare form of breast cancer that begins in the milk ducts and spreads to the skin of the nipple and areola, and only accounts for about 1% of breast cancers. The breast skin may appear crusted, red, or oozing in women suffering from this breast cancer.

Phylloides tumor: Phylloides tumors can either be benign or malignant. They develop in the connective tissues of the breast. Phylloides tumors are very rare.

Scirrhou cancer: This is a breast carcinoma with a very hard texture. Some IDC and LDC may have a scirrhou texture.

APPENDIX B: BDS CLINICAL TRIAL FORMS

B.1 Participant Information Sheet

Version No: 4 Date (24/08/2006)

1. Study Information

Protocol Title:

Early Diagnosis of Breast Cancer with Biopotential Field Technique

Principal Investigator & Contact Details:

Dr. Gregory Kaw Contact: 81263431, Tan Tock Seng Hospital, 11 Jalan Tan Tock Seng, Singapore 308433

Study Sponsor:

Tote-Fund with \$100,000 (for equipment only)

2. Purpose of the Research Study

You are invited to participate in a research study. It is important to us that you first take time to read through and understand this information sheet. Nevertheless, before you take part in this research study, the study will be explained to you and you will be given the chance to ask questions. After you are properly satisfied that you understand this study, and that you wish to take part in the study, you must sign this informed consent form. You will be given a copy of this consent form to take home with you.

You are invited because :

The project's primary objective is to develop a comprehensive clinical diagnostic procedure based on the Biofield Diagnostic System (BDS) and artificial intelligence techniques for the early detection of tumors in the breast. At present the most common methods of screening are by self-examination, clinical examination and mammography. However, it is well known that there can be false negative results with these methods.

Other non-invasive techniques that have great potential for low-cost preliminary screening include techniques based on measuring the bio-field potential and thermal patterns of the breast. The surface biopotential distribution has a relation to the presence, size and location of breast tumors. The surface biofield and temperature distribution can be seen using a general BDS and thermography imager. Both modalities are inexpensive and safe, and could be used for early detection of a breast abnormality. Patients' acceptance towards both methods based on a preliminary study is excellent.

This study aims to recruit **100 breast cancer patients and 30 patients with normal mammograms** from Tan Tock Seng Hospital over a period of 12 months.

3. What procedures will be followed in this study

If you agree to take part in this study, the following will happen to you:

The procedure requires the patient to lie down bare chested (as shown in the accompanying picture) for a few minutes while the surface biopotential field is measured. The entire procedure should be completed in 15 minutes.

There will be no risks/dangers as the BDS process is a painless and harmless diagnostic technique (similar to an electrocardiogram). The bio-field images are acquired from when the scanning probe (with gel) is placed against the skin of the breast. Its many sensors measure the electrical signal at the skin surface (i.e. it detects the small biopotential field from the surface of the breast). This provides functional evidence of breast disease.

You will also be asked to complete the questionnaires. The data collected will be treated confidentially and solely for research purpose only.

No blood will be taken as part of this study.

4. Your Responsibilities in This Study

If you agree to participate in this study, you should follow the advice given to you by the study team. You should be prepared to undergo all the procedures that are outlined above.

5. What Is Not Standard Care or Experimental in This Study

The study is being conducted because BDS is not yet proven to be a standard diagnostic system in subjects with breast cancer. We hope that your participation will help us to determine whether BDS is equal or superior to existing clinical and imaging diagnostic techniques.

In this study, this procedure is only being performed for the purpose of research, and is not part of your routine care.

6. Possible Risks and Side Effects

BDS has no known side effects as it is a breast impedance screening technique. BDS is still being tested; therefore, you may experience other side effects that have not yet been reported. However, you will be kept informed of any significant new findings that may relate to your willingness to continue to take part in this study.

If you experience any new symptoms, you should contact your doctor or the Principal Investigator as soon as possible.

7. Possible Benefits from Participating in the Study

If you participate in this trial you may reasonably expect to benefit from the trial BDS in the following way:

You will have an alternative assessment of your breast condition with the additional BDS screening.

8. Alternatives to Participation

If you choose not to take part in this study, you will still receive standard care for your condition. In our institution this would be mammography and/or ultrasound.

9. Costs & Payments if Participating in the Study

You will not be paid nor you be charged for any study-related procedures while participating in the study.

10. Voluntary Participation

Your participation in this study is voluntary. You may stop participating in this study at any time. Your decision not to take part in this study or to stop your participation will not affect your medical care or any benefits to which you are entitled. If you decide to stop taking part in this study, you should tell the Principal Investigator.

Your doctor, the Investigator and/or the Sponsor of this study may stop your participation in the study at any time if they decide that it is in your best interests. They may also do this if you do not follow instructions required to complete the study adequately. If you have other medical problems or side effects, the doctor and/or nurse will decide if you may continue in the research study.

In the event of any new information becoming available that may be relevant to your willingness to continue in this study, you or your legal representative will be informed in a timely manner by the Principal Investigator or his/her representative.

11. Compensation for Injury

If you follow the directions of the doctors in charge of this study and you are physically injured due to the procedure properly given under the plan for this study, TTSH will pay the medical expenses for the treatment of that injury.

Payment for management of the normally expected consequences of your treatment will not be

Appendix B

provided by TTSH. By signing this consent form, you will not waive any of your legal rights or release the parties involved in this study from liability for negligence.

12. Confidentiality of Study and Medical Records

Information collected for this study including the participants' names and individual results will be kept strictly confidential. Your records, to the extent of the applicable laws and regulations, will not be made publicly available. The data after the completion of the study will remain at TTSH, NP and NTU.

However, the Sponsoring hospital TTSH, Regulatory Agencies and NHG Domain-Specific Review Board and Ministry of Health will be granted direct access to your original medical records to check study procedures and data, without making any of your information public. By signing the Informed Consent Form attached, you or your legal representative is authorizing such access to your study and medical records.

Data collected and entered into the Case Report Forms are the property of TTSH, NP and NTU. In the event of any publication regarding this study, your identity will remain confidential. Different data will be assigned different serial numbers with the key to who the serial numbers belong to kept separately from the data. The procedures to maintain confidentiality and minimize inappropriate data access / release would include: a) secure storage, b) access control, and c) sharing agreements.

13. Who To Contact if You Have Questions

If you have questions about this research study, you may contact the Co- Investigator, Dr Ng Yin Kwee Eddie. Contact: 67904455, School of Mechanical and Aerospace Engineering, Nanyang Technological University.

In case of any injuries during the course of this study, you may contact the Principal Investigator Dr. Gregory Kaw. Contact: 81263431, Tan Tock Seng Hospital, 11 Jalan Tan Tock Seng, Singapore 308433

If you want an independent opinion of your rights as a research subject you may contact the NHG Domain-Specific Research Board Secretariat (Attn: Sujatha Sridhar) at 6471-3266.

A typical BDS sensor arrangement

B.2 Participant Consent Form

Protocol Title:

Early Diagnosis of Breast Cancer with Biopotential Field Technique

Principal Investigator & Contact Details:

Dr. Gregory Kaw Contact: 81263431

Tan Tock Seng Hospital, 11 Jalan Tan Tock Seng Singapore 308433

I voluntarily consent to take part in this research study.
I have fully discussed and understood the purpose and procedures of this study.
This study has been explained to me in _____ (language) on
_____ (date) by _____ (name of translator).

I have been given enough time to ask any questions that I have about the study, and all my questions have been answered to the best of my doctor’s ability.

_____	_____	_____
Name of Patient	Signature	Date

_____	_____	_____
Name of Witness	Signature	Date

Investigator Statement

I, the undersigned, certify to the best of my knowledge that the patient signing this informed consent form had the study fully explained and clearly understands the nature, risks and benefits of her participation in the study.

_____	_____	_____
Name of Investigator	Signature	Date

Appendix B

B.3 Participant Details Form

Participant's Personal Data (for official use)			
NRIC No	Date/Time of Examination Date: _____ Time: _____		
Name	Date of Birth (DD/MM/YYYY)	Weight (kg)	Height (cm)
Nationality	Race	Marital Status	Age at first live birth
Smoking: Y ☐ / N ☐	Alcohol: Y ☐ / N ☐		
Age at Menarche:	Average no. of days between menstrual cycles	Menopause Y ☐ / N ☐ If yes, what age? _____ &	
No. of days since the first day of last menstrual cycle:	Medicines taken	Hormone Replacement Treatment Y ☐ / N ☐	
Any previous history of breast cancer? Y ☐ Year: _____ N ☐			
Any relatives with breast cancer? Y ☐ / N ☐ If yes, Relationship _____ Age diagnosed _____			
Tick Yes No _____ Pain? Since _____ If Yes Circle: Right Left _____ Lumps in either breast? Since _____ If Yes Circle: Right Left _____ Discharge from Nipple? If Yes Circle: Right Left _____ Is the discharge bloody? If Yes Circle: Right Left _____ Any breast Surgery or breast biopsy? If Yes, Year _____ Circle: Right Left Biopsy results _____ _____ Previous Mammogram? When _____ If Yes, Where _____			
Bra Cup Size _____			

PARTICIPANT FEEDBACK FORM

- Did you feel any pain during the test? Yes ☐ No ☐
- Is there any itching where the sensors were placed? Yes ☐ No ☐
- How will you compare the comfort provided by BDS with the comfort of mammogram and ultrasound?
More comfortable ☐ Similar ☐ Less comfortable ☐
- Are you satisfied with the speed of BDS compared to speed of Ultrasound and Mammogram? Yes ☐ No ☐
- Would you undergo this test again? Yes ☐ No ☐
- Would you recommend this test to your friends and family? Yes ☐ No ☐

Appendix B

Diagnosis (for official use)		
Clinical Breast Examination Results		
Palpable? Y ☐ / N ☐ If yes, indicate on diagram	Right Breast	Left Breast

Ultrasound Results (if available):		Date:
Result:		LOS:
Tumor size:		
Mammogram Results (if available):		Date:
Result:		LOS:
Breast density:		
Biopsy Results (if available):		Date:
BDS Results		Date:
LOS (according to clinical diagnosis by doctor):		1 2 3
BDS index:		
BDS result:		
Filled by:		
Signature/Date:		

B.4 BDS Sample Test Result Form

Diagnostic System Test Result Form

Patient Name	<input type="text"/>	Test Date	<input type="text"/>
Patient Number	<input type="text"/>	Test Time	<input type="text"/>
Date of Birth	<input type="text"/>	Tested Lesion	<input type="text"/>
Physician Name	<input type="text"/>		
Operator Name	<input type="text"/>		
Signature of Person Performing the Test _____		Date _____	
		<div>Right Left</div> <div>Indicate Location of Tested Lesion (X)</div>	
		Indicate Breast Map Used: _____	
<div><div></div></div>			

Appendix B

Based on the results of clinical trials in women aged 18 to 55 with palpable lesions, the following chart shows how to combine the BDS index with a prior level of suspicion to obtain an adjusted level of suspicion.

The BDS index should **NOT** be used to adjust prior level of suspicions of 4 ("Probably Malignant") or 5 ("Malignant").

PRIOR LEVEL OF SUSPICION

Clinician should circle the prior level of suspicion below:

BDS Index:	Level of Suspicion: 1	Level of Suspicion: 2	Level of Suspicion: 3
	Normal/ No Finding	Probably Benign	Malignancy Not Excluded
-5.0 to -1.7	Probably Benign	Probably Benign	Probably Benign
-1.6 to +0.2	Probably Benign	Probably Benign	Malignancy Not Excluded
+0.3 to +1.5	Malignancy Not Excluded	Probably Benign	Probably Malignant
+1.6 to +5.0	Malignancy Not Excluded	Malignancy Not Excluded	Probably Malignant

Prior "Level of Suspicion" is determined from physical or imaging examination and is defined as follows:

- 1. Normal/No Finding
- 2. Probably Benign
- 3. Malignancy Not Excluded

These results must be interpreted using additional information in the Biofield Diagnostic System Physician's Manual (Part Number 725-10004-00X).

Notes: