1-1-2000

Multiattribute Decision-Making: Use of Three Scoring Methods to Compare the Performance of Imaging Techniques for Breast Cancer Detection

Fred S. Azar
University of Pennsylvania

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-00-10.

Multiattribute Decision-Making: Use of Three Scoring Methods to Compare the Performance of Imaging Techniques for Breast Cancer Detection

Abstract

Multiple Attribute Decision Making (MADM) involves "making preference decisions (such as evaluation, prioritization, selection) over the available alternatives that are characterized by multiple, usually conflicting, attributes". The problems of MADM are diverse, and can be found in virtually any topic.

In this paper, we use three different scoring methods for evaluating the performance of different imaging techniques used to detect cancers in the female breast. The need for such a decision support system arises from the fact that each of the several techniques which helps diagnose breast cancer today, has its own specific characteristics, advantages and drawbacks. These characteristics or attributes are generally conflicting.

The goal is to detect as many malignant lesions in the breast as is possible, while identifying the maximum number of benign lesions. The four imaging techniques that are compared here are Magnetic Resonance Imaging (MRI), Mammography, Ultrasonography, and Nuclear Medicine.

The three different multiattribute scoring methods are the Simple Additive Weighting method (SAW), the Weighted Product Method (WPM), and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). The three methods are described in detail, and then used to rank the four imaging techniques. The results are analyzed and the validity and robustness of the methods are tested using post-evaluation analysis.

Disciplines

Oncology

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-00-10.

Multiattribute Decision-Making: Use of Three Scoring Methods to Compare the Performance of Imaging Techniques for Breast Cancer Detection

Fred S. AZAR

University of Pennsylvania, Philadelphia (PA)
Dept. of BioEngineering
VAST LAB, Dept. of Computer Science

Technical Report MS-BE-00-01
MS-CIS-00-10
2000

1 INTRODUCTION

Multiple Attribute Decision Making (MADM) involves “making preference decisions (such as evaluation, prioritization, selection) over the available alternatives that are characterized by multiple, usually conflicting, attributes” [9]. The problems of MADM are diverse, and can be found in virtually any topic. Even Ben Franklin, over 200 years ago recognized the presence of multiple attributes in everyday decisions, and suggested a workable solution [10].

In this paper, we use three different scoring methods for evaluating the performance of different imaging techniques used to detect cancers in the female breast.

The need for such a decision support system arises from the fact that each of the several techniques which helps diagnose breast cancer today, has its own specific characteristics, advantages and drawbacks. These characteristics or attributes are generally conflicting.

The goal is to detect as many malignant lesions in the breast as is possible, while identifying the maximum number of benign lesions. The four imaging techniques that are compared here are Magnetic Resonance Imaging (MRI), Mammography, Ultrasonography, and Nuclear Medicine.

The three different multiattribute scoring methods are the Simple Additive Weighting method (SAW), the Weighted Product Method (WPM), and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS).

The three methods are described in detail, and then used to rank the four imaging techniques.

The results are analyzed and the validity and robustness of the methods are tested using post-evaluation analysis.

2 CHOOSING THE RIGHT IMAGING TECHNIQUE FOR BREAST CANCER DETECTION

The emphasis on early detection of breast cancer, the desire not to miss a malignant lesion in the early stage of disease, and the current medicolegal environment encourage an aggressive biopsy approach to breast problems [8]. With such an approach, a large majority of the palpable and mammographically detected nonpalpable breast lesions on which biopsies are performed will be benign [8]. Actually the positive biopsy rate for cancer is low (10%-31% [15][16][17]), meaning that 70%-90% of breast biopsies are performed in women with benign disease [8].

Biopsies, although well tolerated, involve some risk, induce patient discomfort and anxiety, and increase costs in terms of both patient recovery and overall health care expenses [8]. Indeed, according to Stavros and al. [8], the total cost for percutaneous large-core biopsy (including the technical component, the physician's fee, and the pathology fees) of a breast nodule with ultrasound guidance is approximately $1000. The total cost for an excisional breast biopsy is between $3000 and $4500.

It is therefore imperative to choose the best imaging technology in order to reduce the negative-to-positive biopsy ratio, and therefore the cost to society (including health care expenses). The financial savings could be considerable. Furthermore, the morbidity associated with the biopsy procedure including the lost time from work that occurs, as a result of biopsy could be greatly reduced [8].

3 METHODS FOR MULTIATTRIBUTE DECISION MAKING

Hwang and Yoon classified a group of 17 multiattribute decision-making methods according to the type and important features of information received [9].

We have selected three methods that were found well adapted to the decision at hand. These methods enable us to obtain a meaningful **utility index** from multidimensional data to evaluate the competing alternatives.

The three methods are Simple Additive Weighting (SAW), Weighted Product Method (WPM), and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). They are described in detail in the following sections.

The analysis of the problem is carried out, generally following the systematic procedure described in [1], which is an augmented version of the Simple Multiattribute Rating Technique (SMART) from Ward Edwards [19].

All three methods share a common number of steps:

- Determine the relevant goals
- Determine the Alternatives to be evaluated
- Identify the relevant attributes for evaluating the alternatives, and determine the values over which the individual attributes range
- Rank and obtain weights (which are normalized) for the attributes in order of relative importance
- Score each attribute of each alternative. Each alternative i may be thought of as a vector of descriptions on separate attributes j [1]: $X_i = < x_{i,1}; x_{i,2}; \ldots; x_{i,j}; \ldots; x_{i,n} >$
- Calculate the multiattribute utilities of the various options, using the three scoring methods
- Perform post-evaluation analysis, and then decide which is the best alternative

It is important to note that all attribute ratings are normalized in a specific way using each different method, to eliminate computational problems caused by differing measurement units. Normalization aims at obtaining comparable scales, which allow interattribute as well as intra-attribute comparisons [2]. Therefore normalized ratings have dimensionless units and, the larger the rating becomes, the more preference it has [1].

4 ATTRIBUTE GENERATION, DATA, AND WEIGHT

4.1 What Goals are Relevant?

The "super goal" is to determine the imaging technique that is best suited for correctly diagnosing breast cancer. The user, who is a doctor or radiologist, needs the technique that will help him/her identify all of the malignant tumors in a breast, while recognizing the other tumors, which are benign.

4.2 Alternatives to be evaluated

The different techniques involved in this study are:

- Magnetic Resonance Technology
  A large magnetic field created by an MR machine is used to image the internal structures of the breast. Images obtained usually have one of the highest resolutions available today.
- Mammography
  Here the X-rays generate the image on a film, after traversing the breast.
- Ultrasonography
  Ultrasounds are 'bounced' against the structures in the breast and the waves that bounce back are responsible for generating the image of the breast. Ultrasound has a relatively high resolution however, the images obtained are usually very hard to interpret.
- Nuclear Medicine
  Positron Emission Tomography (PET), a technique used in nuclear medicine, can be used to assess the location and nature of cancer lesion by determining the moment-to-moment changes in the chemistry or physiology of compounds that are administered to the patient and labeled with a radioactive isotope. The image is the distribution of radioactivity in the breast of the patient.

The alternatives considered here constitute the main techniques used today which help the diagnosis of the physician.

There actually are several subtechniques for each of those alternatives, and sometimes two techniques are used concurrently:

Nuclear medicine breast imaging techniques with use of positron-emitting radiotracers are being considered as complementary modalities to conventional mammography [4], because they can obtain functional or metabolic information that could potentially decrease the number of unnecessary biopsies performed [4] [7].

Artificial neural networks have been used along with mammography to increase the ability to discriminate between benign and malignant lesions [5]. Other computer programs have been used which issue an advisory estimate of the probability of malignancy [6].

However considering all of them would be a very hard task, since the information needed to evaluate them is not readily available.

In comparing the four main technologies, we have considered the best attainable results in each of these technologies, in order to be fair, and to obtain meaningful results.

4.3 Identifying the Relevant Attributes

The attributes we are seeking should be the most important ones deemed relevant to the final decision. They should preferably be mutually exclusive: the attributes should be viewed as independent entities among which appropriate trade-offs may later be made [2].

Most importantly, the chosen attributes should be measurable in a meaningful and practical way, for each of the proposed alternatives.

When considering the attributes for imaging techniques, one might first think of such quantities as ‘resolution’ or ‘definition’. However those dimensions are too general, and do not accurately reflect the performance of the four imaging techniques when it comes to breast imaging, and the goals we have in mind.

We might then consider such attributes as the ability to distinguish relevant features in the image, such as the shape of the tumor, its size, its aspect, the difference in contrast between the inside and the outside of the tumor. These attributes are legitimate when considering a technique such as *MR imaging*.

Detectability of a cancer using a technique such as scintimammography (*nuclear medicine*), is determined by radiotracer uptake in the tumor, attenuation of gamma rays from the tumor by intervening tissue, intrinsic detector spatial resolution, separation between the tumor and the camera.

Radiographic features used in the interpretation of *mammograms* can be categorized into those related to masses (shape, size, pattern), to microcalcifications (number, shape, distribution), and to secondary abnormalities (parenchymal distortion, skin thickening) [5].

When considering *ultrasonography*, reported relevant characteristics include shadowing, hypoechogenicity, microlobulation, or ellipsoidal shape [8].

We have found that the only way that is efficient and fair, and which correlates directly to the goals of the study, is to evaluate the techniques based on their ability to actually detect the tumors from reported experiments. Therefore, such attributes as Sensitivity or Specificity will be much more adapted to the task at hand, and are readily measurable quantities.

The first four attributes are directly related to the performance of the imaging techniques. They are determined after studies in which several physicians have used the techniques to actually classify malignant and benign lesions in the breasts of many patients.

It is useful to define the following measured quantities first [3]:
TP = True Positives: Nb. of malignant findings correctly classified
TN = True Negatives: Nb. of benign findings correctly classified
FP = False Positives: Nb. of benign findings incorrectly classified
FN = False Negatives: Nb. of malignant findings incorrectly classified

Then the first four measurable attributes are (as explained in[3]):

1) Sensitivity (SE) = TP / (TP + FN)
2) Specificity (SE) = TN / (TN + FP)
3) Positive Predictive Value (PPV) = TP / (TP + FP)
4) Negative Predictive Value (NPV) = TN / (TN + FN)

These first four attributes are naturally measured in percentage, and therefore range from 0 to 100.

The following two attributes which were proposed by several doctors and/or professors in the hospital of the University of Pennsylvania (HUP, radiology section), relate to the complexity of the imaging technique: complexity with respect to the physician, and complexity with respect to the patient:

5) Complexity of Interaction with Patient (CIP)
   (Includes time spent, degree of comfort, invasiveness...)
6) Complexity of Interaction with Doctor (CID)
   (Includes time spent, level of necessary training and experience, complexity of protocol...)

Attributes 5) and 6) have been given a Likert-type scale [20][21], since they are qualitative data to be assigned numerical values. An arbitrary integer range from 1 to 10 was created as follows:

- 1-2 : very simple
- 3-4: simple
- 5-6: not too complex
- 7-8: complex
- 9-10: very complex

Radiologists and/or physicians have been asked to rate each of the four imaging technologies from 1 to 10.

The last attribute, which is as relevant to the patient as it is to the physician, is the cost in dollars per operation (to the patient and/or the insurance company), for each of the imaging techniques:

7) Cost ($) per operation (C)

The cost per operation varies in a range from $50 (a cheap mammography) to $700 (an expensive MR exam).

4.4 Ranking the Attributes and Obtaining relative weights

The Attributes were ranked following the priority number one: Correctly identify the maximum number of malignant lesions in a given patient breast.

We also assumed that the Cost of operation for each machine comes as priority number two, and that the complexity of interaction with the physician and patient is not as relevant as the cost.

The final ranking of the attributes, as well as the weights w (which were subjectively chosen after consulting a physician in the radiology dept) are given as follows, most important first:

1) SE Nominal Weight: w=55
2) PPV Nominal Weight: w=45
3) SP Nominal Weight: w=40
4) NPV Nominal Weight: w=35
5) C Nominal Weight: w=30
6) CIP Nominal Weight: w=20
7) CID Nominal Weight: w=15

4.5 Scoring each attribute of each alternative imaging technique

The scores for SE, PPV, SP and NPV were determined as follows:

a) Mammography:

- We obtained the scores from a study that compared the performances of a neural network and of radiologists in the task of distinguishing between benign and malignant breast lesions [5]. The database consisted of features from 133 textbook cases and 60 clinical cases.
- Other sources show that mammography and physical examination combined have a sensitivity of 85% for the detection of breast carcinoma. Mammography also has a PPV of 15%-30% [7][14].

b) Nuclear medicine:

- A recent European multicenter trial that evaluated palpable and nonpalpable breast lesions using scintimammography demonstrated an overall sensitivity and specificity of 80% and 73% respectively [11].
- Scintimammography has excellent sensitivity for tumors larger than 1cm, but is generally poor for smaller, or nonpalpable lesions [4]: A recent three-center study in Europe reported sensitivities of 26% (<0.5cm), 56% (0.5-1.0cm), 95% (1.2cm), and 97% (>2.0cm) [12].
- Two prospective multicenter studies have demonstrated a negative predictive value of 94% for palpable lesions [18].
- Conventional whole body Positron Emission Tomography has reported sensitivities and specificities that range between 70% to 90% and 85% to 95%, respectively [4].

c) MRI:

- The scores obtained come from a study of 192 patients with mammographically visible or palpable findings. Patients underwent subsequent excisional biopsy for histopathologic confirmation. An interpretation model was constructed by using 98 cases and was tested prospectively and expanded by using 94 different cases [3].
- Although very sensitive for breast cancer, MR imaging does suffer from a relatively low specificity [13].

d) Ultrasonography:

- The data used comes from a study performed on 750 sonographically solid breast nodules, which were classified as benign, indeterminate or malignant [8]. The

classifications were compared with biopsy results, and SE, PPV, SP, and NPV were computed.
- The scores we used are the averages of each attribute over the nine most important reported characteristics found in Table 6 of [8].

Each of the scores for C, CIP and CID was obtained after consulting several physicians/radiologists, at the hospital of Pennsylvania.

The different scores are given as follows:

Options		SE	PPV	SP	NPV	C	CIP	CID
S1: ***MRI Tech***	Scores:	96	75	69	97	500	9	5
S2: ***Mammography***	Scores:	89	57	45	83	75	2	2
S3: ***Ultrasound***	Scores:	48	65	94	90	200	6	9
S4: ***Nuclear Medicine***	Scores:	80	41	73	94	350	4	7

<u>Note</u>: the NPV for mammography and the PPV for nuclear medicine were computed using the following derived equations:

$$NPV = \frac{SP}{SP + \frac{(1-SE)\cdot(1-SP)\cdot PPV}{(1-PPV)\cdot SE}} \quad \text{And} \quad PPV = \frac{SE}{SE + \frac{(1-SP)\cdot(1-SE)\cdot NPV}{(1-NPV)\cdot SP}}$$

5 FIRST MODEL : SIMPLE ADDITIVE WEIGHTING METHOD (SAW)

The SAW method is probably the best known and most widely used. The total score for each alternative is computed by multiplying the comparable rating (or utility) for each attribute by the importance weight assigned to the attribute and then summing these products over all the attributes.

5.1 The Unidimensional Utility Functions

The unidimensional utility (or value) functions u_i have the mathematical form [1]:

$$u_i(x_{ij}) = 100 \times \left(\frac{(x_{ij} - c_i)}{(b_i - c_i)} \right)^{r_i}$$

where u_i is the utility function for attribute i , where i belongs to {SE, PPV, SP, NPV, C, CIP, CID}

x_{ij} is the score of attribute i for alternative j , where j belongs to {MRI, Ultrasound, Mammography, Nuclear Medicine}

r_i is the risk aversion factor (utility is risk neutral if r=1, risk averse if 0<r<1, risk seeking if r>1)

b_i and c_i are the values of best and worst outcome respectively, for attribute i. They constitute the range of the attribute.

5.2 Determining the risk aversion factors

Assumption: $\underline{u_i\,(x_{ij\text{-}lowest}) = 50}$
i.e. **The utility of the lowest score, for each attribute is 50**.

This assumption comes from the fact that the four available imaging techniques are used today, in different circumstances, but all of them are useful, and do score more than average i.e. 50%, on average, since they are in use by thousands of physicians today!

Of course, this is a subjective decision, and other schemes are possible, however this assumption seems reasonable enough.

Then we can find the risk aversion factor from the following computed equation:

$$r_i = \frac{\ln\left(\dfrac{50}{100}\right)}{\ln\left(\dfrac{x_{ij-lowest} - c_i}{b_i - c_i}\right)} = \frac{\ln(2)}{\ln(b_i - c_i) - \ln(x_{ij-lowest} - c_i)}$$

We obtain:

1) $r_{SE} = 0.9444$
2) $r_{PPV} = 0.7774$
3) $r_{SP} = 0.8680$
4) $r_{NPV} = 3.720$
5) $r_C = 0.5880$
6) $r_{CIP} = 0.3155$
7) $r_{CID} = 0.3155$

5.3 The MultiAttribute Utility Function

The multiattribute utility function U is a linear combination of the u_i's. $U(X_j)$ is the final utility or value of alternative X_j.

$$U(X_j) = \sum_{i=1}^{7} w_i \cdot u_i(x_{ij})$$

where

$$X_j = \langle x_{1j} \quad x_{2j} \quad x_{3j} \quad x_{4j} \quad x_{5j} \quad x_{6j} \quad x_{7j} \rangle$$

6 SECOND MODEL : WEIGHTED PRODUCT METHOD (WPM)

The Weighted Product Method was introduced by Bridgeman [22]. According to [2], the method possesses sound logic, and is computationally simple, but has not been widely utilized.

Contrary to the SAW method, the different measurement units here do not have to be transformed into a dimensionless scale by a normalization process. This is because in the WPM method, the attributes are connected by multiplication.

The weights become exponents associated with each attribute value (positive power for benefit attributes, and negative power for cost attributes) [2].

The multiattribute utility function U of alternative X_j is given by:

$$U(X_j) = \frac{\prod_{i=1}^{7} x_{ij}^{w_i}}{\prod_{i=1}^{7} \left(x_i^*\right)^{w_i}}$$

Where x_i^* is the most favorable value (i.e. The best score among the four alternatives) for the i^{th} attribute, and belongs to X^*, the “ideal alternative”.

The use of X^* allows us to put a numerical upperbound to the alternative values obtained by this multiplicative method. Hence, by comparing each alternative with the ideal alternative we can see that U is here between 0 and 1.

It is important to note that this method requires that all scores be greater than 1, because of the exponent property (which the case for our study).

7 THIRD MODEL : TECHNIQUE FOR ORDER PREFERENCE BY SIMILARITY TO IDEAL SOLUTION (TOPSIS)

Hwang and Yoon [9] developed the TOPSIS technique based on the concept that “the chosen alternative should have the shortest distance from the positive–ideal solution and the longest distance from the negative-ideal solution” [2]. The ideal solution is the collection of ideal scores (or ratings) in all attributes considered.

The TOPSIS technique defines a “similarity index” (or relative closeness) by combining the proximity to the positive-ideal solution and the remoteness of the negative-ideal solution [2].

Several steps are needed in order to implement the technique [2]:

7.1 Calculate Normalized Scores

Vector normalization is used:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^{4} x_{ij}^2}}$$ where j is the index related to the alternatives, and i to the attributes.

7.2 Calculate Weighted Normalized Ratings

The weighted normalized value is calculated as:
$v_{ij} = w_i \cdot r_{ij}$ where w_i is the weight of the ith attribute.

7.3 Identify Positive-Ideal and Negative-Ideal Solutions

The positive-ideal solution is the composite of all best attribute ratings attainable, and is denoted:
$A^* = \{v_1^*, v_2^*, ..., v_i^*, ..., v_n^*\}$ where v_i* is the best value for the ith attribute among all alternatives.

The negative-ideal solution is the composite of all worst attribute ratings attainable, and is denoted:
$A^- = \{v_1^-, v_2^-, ..., v_i^-, ..., v_n^*\}$ where v_i^- is the worst value for the ith attribute among all alternatives.

7.4 Calculate Separation Measures

The separation or distance of each alternative from the positive-ideal solution A^*, is given by the n-dimensional Euclidean distance:

$$S_j^* = \sqrt{\sum_{i=1}^{7} (v_{ij} - v_i^*)^2}$$ where j is the index related to the alternatives, and i to the attributes.

Similarly, the separation from the negative-ideal solution A^- , is given by:

$$S_j^- = \sqrt{\sum_{i=1}^{7} (v_{ij} - v_i^-)^2}$$

7.5 Calculate Similarity Indexes

The Similarity to positive-ideal solution, for alternative j, is finally given by:

$$C_j^* = \frac{S_j^-}{S_j^* - S_j^-} \quad \text{with} \quad 0 \le C_j^* \le 1$$

The alternatives can then be ranked according to C_j^* in descending order.

8 IMPLEMENTATION OF THE SCORING METHODS

The implementation of the different methods, and the analysis that follows were done using Microsoft Excel97.

Indeed a spreadsheet program such as Excel offered all the necessary and desirable tools to complete the analysis of the multiattribute decision methods. We made use of the layout, the formatting of objects, the absolute and relative addressing, the formulas, the charting capabilities, the data tables and the goal seeking features.

We also made use of Visual Basic to create functions, such as the unidimensional utility functions used in the SAW method.

The Parameters Page is shown below, and contains all user-defined parameters (the input parameters are shaded).

Name	Value	Comment
w1nom	55	nominal value for w1
w2nom	45	nominal value for w2
w3nom	40	nominal value for w3
w4nom	35	nominal value for w4
w5nom	30	nominal value for w5
w6nom	20	nominal value for w6
w7nom	15	nominal value for w7
Risk1	0.9444	risk coeff for (SE)
Risk2	0.7774	risk coeff for (PPV)
Risk3	0.868	risk coeff for (SP)
Risk4	3.72	risk coeff for (NPV)
Risk5	0.588	risk coeff for (C)
Risk6	0.31546	risk coeff for (CIP)
Risk7	0.31546	risk coeff for (CID)
x1hi	100	best outcome on (SE)
x1lo	0	worst outcome on (SE)
x2hi	100	best outcome on (PPV)
x2lo	0	worst outcome on (PPV)
x3hi	100	best outcome on (SP)
x3lo	0	worst outcome on (SP)
x4hi	100	best outcome on (NPV)
x4lo	0	worst outcome on (NPV)
x5hi	50	best outcome on (C)
x5lo	700	worst outcome on (C)
x6hi	1	best outcome on (CIP)
x6lo	10	worst outcome on (CIP)
x7hi	1	best outcome on (CID)
x7lo	10	worst outcome on (CID)

Name	Value	Comment
w1ini	55	inital value for w1
w2ini	45	inital value for w2
w3ini	40	initial value for w3
w4ini	35	initial value for w4
w5ini	30	initial value for w5
w6ini	20	initial value for w6
w7ini	15	initial value for w7

ratios	
ratio1	1.222222
ratio2	1.125
ratio3	1.142857
ratio4	1.166667
ratio5	1.5
ratio6	1.333333

kratio	1	Value of kratio
kratioMAX	1.125	(in order to preserve order of priority)
kratioMIN=0.1		

INCLUDE COST IN STUDY ? (yes=1, no=0) : 1

9 RESULTS

The initial results are as follows:

Alternatives	*Scoring Methods*						Average of Methods	
	SAW		WPM		TOPSIS			
	Rank	Score	Rank	Score	Rank	Score	Rank	Score
S1: ***MR Imaging***	1	77.750	2	62.420	3	47.932	2	62.701
S2: ***Mammography***	2	74.532	1	80.710	1	65.039	1	73.427
S3: ***Ultrasonography***	4	70.812	3	60.365	2	51.339	3	60.839
S4: ***Nuclear Medicine***	3	72.643	4	58.840	4	45.907	4	59.130

The initial rankings don't seem to agree too well, with SAW ranking MR imaging on the top, but TOPSIS ranking it only in third position. Mammography gets first or second positions in all methods.

However, if we consider the average of the three methods, then, Mammography is first, followed by MRI. The scores for Ultrasonography and Nuclear Medicine are extremely close, which puts them both roughly in third position.

Comparing these results with today's fact, we know that mammography is the technique most widely used for breast cancer detection. MRI is being also used, but hasn't been widely adopted. Whereas Ultrasonography and Nuclear Medicine are being used today in conjunction with Mammography and MRI, in order to help with the diagnosis (however this is still not a wide practice).

According to most physicians, The reason why mammography is used more widely for breast cancer detection is that it is a cheap technique to use, and most health institutions can afford to buy mammographic equipment.

How would the results then vary, if cost C was not part of the equation, and was ignored? The answer to this question follows:

Alternatives	*Scoring Methods*						Average of Methods	
	SAW		WPM		TOPSIS			
	Rank	Score	Rank	Score	Rank	Score	Rank	Score
S1: ***MR Imaging***	1	81.714	2	76.522	1	61.096	1	73.111
S2: ***Mammography***	3	71.220	1	78.276	2	57.601	2	69.032
S3: ***Ultrasonography***	4	68.685	4	64.613	4	45.965	4	59.754
S4: ***Nuclear Medicine***	2	73.094	3	67.973	3	49.692	3	63.586

The above scores and rankings show that if the cost C is not a major concern, then MRI becomes the winner, followed by mammography, nuclear medicine, and ultrasonography.

Indeed, MRI imaging seems to score better in the major attributes that are directly related with the ability to detect a malignant cancerous lesion in the breast. HOWEVER, an MR machine is much more expensive to operate, and requires special attention, and maintenance (this is reflected in the high cost on the patient, per exam).

The results agree with the facts: An area of technical development is now in the field of low-cost, dedicated breast MR systems. Despite its potential diagnostic benefits, clinical application of breast MR imaging may become a casualty of its high cost [4]. High utilization of low-cost, dedicated systems could reduce the cost of breast MR imaging dramatically.

The other fact is that there is an emergence of techniques using a combination of technologies, such as mammography and nuclear medicine on one side, and MR and ultrasonography on the other. The results do not contradict that fact.

10 SENSITIVITY ANALYSIS

Before we can safely draw a conclusion from the results, we must ask ourselves the following question. How do the user defined parameters in each of the three scoring methods influence the final score, and hence the final decision?

This is where sensitivity analysis kicks in. By varying one or more parameters, we can study the variations of the results, and thus examine the 'robustness' of the final decision.

Among the different parameters, the most subjective ones must be the weights assigned to each attribute. Therefore, by varying the different weights we will be able to study their effect on the final scores. It seems that the cost C had a strong effect on the final decision. Furthermore, the high value of the weight assigned to Sensitivity SE shows that it must have a strong influence on the overall scores.

Note: Sensitivity analysis is where we can appreciate the ease of use and power of an EXCEL spreadsheet implementation of the problem.

10.1 Effect of w_{SE} on Final Scores

Creating data tables enabled us to display interesting charts regarding the variation of the various utilities. The effect of w_{SE} on the final scores is shown in Appendix A1

For SAW, we see clearly that MRI and Mammography remain first and second respectively for all values of W_{SE}. Whereas ultrasonography decreases sharply, due to its low score on Sensitivity.

For WPM, it is mammography that is largely in first position, and it would take meaningless large value of W_{SE} for MRI to capture first position

For TOPSIS, Mammography still is first, but MRI rises and reaches 2^{nd} position at $W_{SE} = 62$.

Overall, when considering the average result, Mammography wins, and MRI is second starting at W_{SE}=50. Nuclear medicine rises to be third, whereas ultrasonography decreases sharply.

10.2 Effect of w_C on Final Scores

The effect of W_C on the final scores is shown in Appendix A2.

The result for SAW tells us that if we give enough importance to the Cost attribute (as much importance as sensitivity), then mammography replaces MRI for first position.

The result for WPM clearly shows two tendencies, the utility of mammography goes up with w_c and the utilities of all other alternatives drop clearly much lower than the utility of mammography. This behavior is probably due to the mathematical formulation of WPM, and the fact that the weights are exponents. This result does not seem to show any kind of meaning.

We see that, using TOPSIS, MRI is the best alternative as long as the cost attribute remains the least important (w_c<10). Beyond that point, mammography takes the lead.

The result of the average of the three techniques shows that it is mostly influenced by the results for WPM, hence we cannot derive much meaning from it.

10.3 Effect of Weight Ratio Variation

We introduce here the concept of Weight Ratio Variation (WRV).
We want to vary the distribution of the relative ratios of the nominal weights.
Let the existing ratios be defined as:
$w_{SE} / w_{PPV} = r_1$, $w_{PPV} / w_{SP} = r_2$ … and so on.

We create new ratios r'_i such that: $\mathbf{r'_i = r_i / k_{ratio}}$

Then starting from the least important attribute (that we leave unchanged), we recalculate the nominal weights, satisfying those new ratios:

- $w'_{CID} = w_{CID}$
- $w'_{CIP} = r'_6 \, . \, w'_{CID}$
- $\mathbf{w'_i = r'_i \, . \, w'_{i-1}}$

When k_{ratio} is 1, the new and old weights are equal. When $k_{ratio} > 1$, the relative ratios of the nominal weights are smaller, and the weights are distributed more closely to each other. When $k_{ratio} < 1$, the weights are distributed farther away from each other.

In general, if we wish to keep the initial ranking of the attributes unchanged, then we should make sure that: $k_{ratio\text{-}MAX} < r_{i\text{-}MIN}$

Otherwise, by increasing k_{ratio} beyond that value, the ranking of the attributes will gradually reverse itself, until $k_{ratio} = r_{i\text{-}MAX}$, value beyond which the attributes' ranking will be completely reversed.

It is interesting to study the effect of the variation of k_{ratio} on the final scores of each method. Using this variation, we can see how robust is the initial choice of the relative importance of an attribute with respect to the other, and therefore verify the validity of the initial choice.

The results are shown in Appendix A3.

At the end of Appendix A3 is shown a graph of the relative weight factors with respect to the initially lowest weight $w_{CID} = 15$. This graph will help make sense of the other graphs, since it shows us the relative importance of weights with respect to each other.

The results for SAW are interesting and show that k_{ratio} can be increased up to 1.12 before mammography takes over MRI for first position. This shows the robustness of the method, with the chosen initial parameters.

The results for WPM still, as in the preceding section, don't seem to hold much meaning, except for the fact that mammography is a clear winner in all situations! Again, it is the properties of the mathematical formulation of the method which are responsible.

The results for TOPSIS show that MRI is the leading choice if $k_{ratio} < 0.86$, that is, if sensitivity SE is 9 times more important than CID, and 3.33 times more important than cost C. However, for $k_{ratio} > 0.86$, mammography becomes the clear winner.

11 CONCLUSION

Three multiattribute decision making methods have been implemented using MS excel97, to determine which of four different imaging technologies is best suited for achieving the best results in breast cancer detection.

MRI, Mammography, Ultrasound, and Nuclear Medicine were compared to each other, using experimental data and human studies found in recent medical literature, and using the expertise of a few doctors and professors in the dept of Radiology of the Hospital of the University of Pennsylvania.

After analysis, and post-evaluation analysis, using graphs from data tables, we concluded that this evaluation does indeed make a lot of sense, and agrees with the existing facts.

We found that the most robust method seems to be the SAW method, which explains its popularity.

The TOPSIS method yielded results similar to SAW but it appears that the cost factor C had more influence on the final ranking, clearly benefiting mammography and ultrasonography because of their lower operating costs.

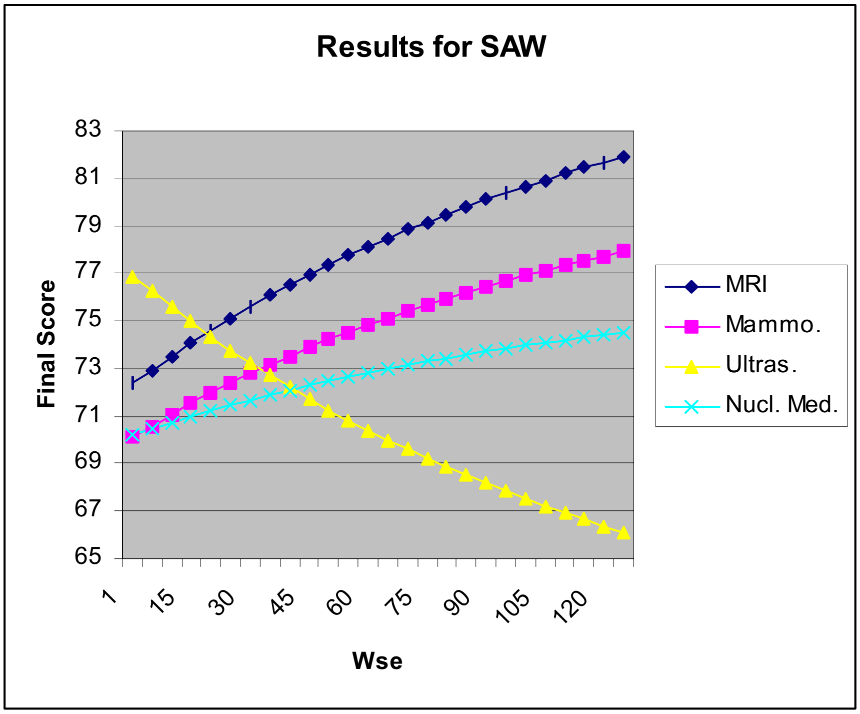
The WPM method yielded results that were too extreme to be taken into account. One has to be careful when using this method, because of the use of weights as exponents in the mathematical formulation. The behavior of exponential functions is probably responsible for the extreme results obtained.

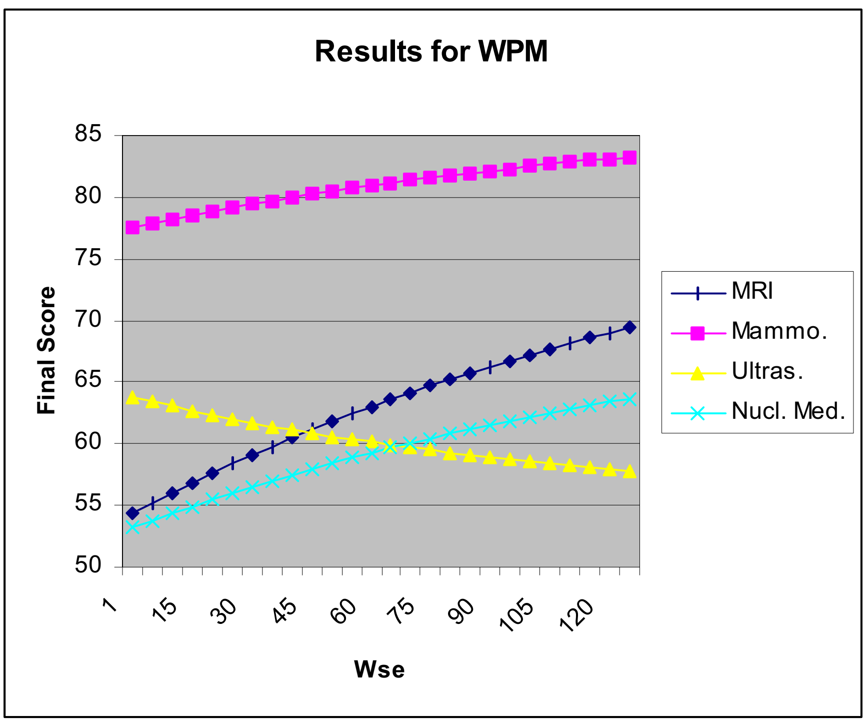
The analysis concludes that MRI is the best technology for classifying malignant cancer lesions in the breast, and should be used **if Cost is not an issue**. MR imaging may play an important future role in guiding therapy for breast cancer. MR imaging already provides an extremely accurate map of the extent of cancer within an affected breast [4].

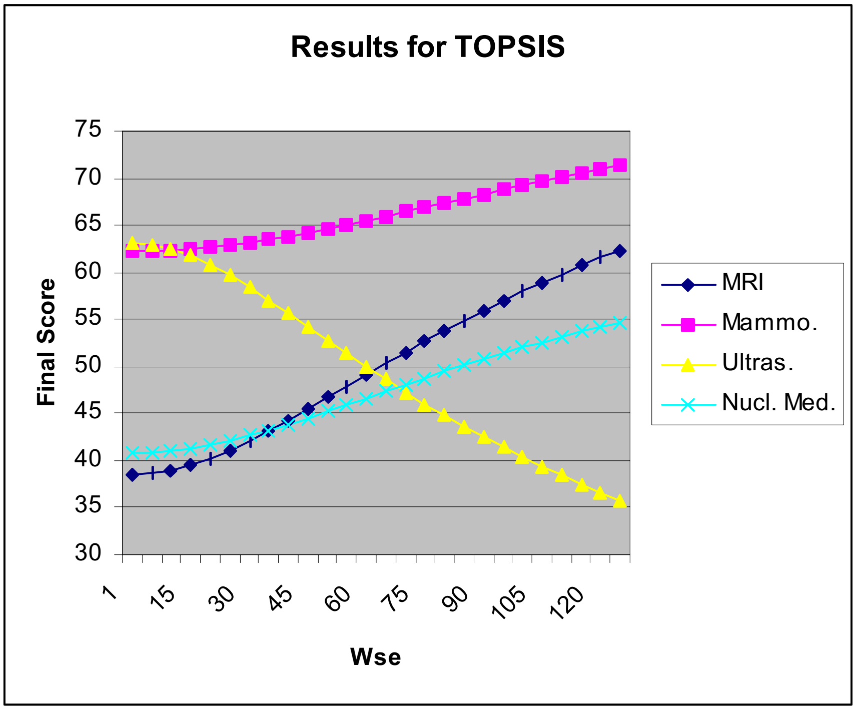
If Cost is the leading factor, or second leading factor in the choice of technology, then <u>mammography remains the leader</u>, which explains why so many hospital facilities still use mammography, even though MRI yields better results.

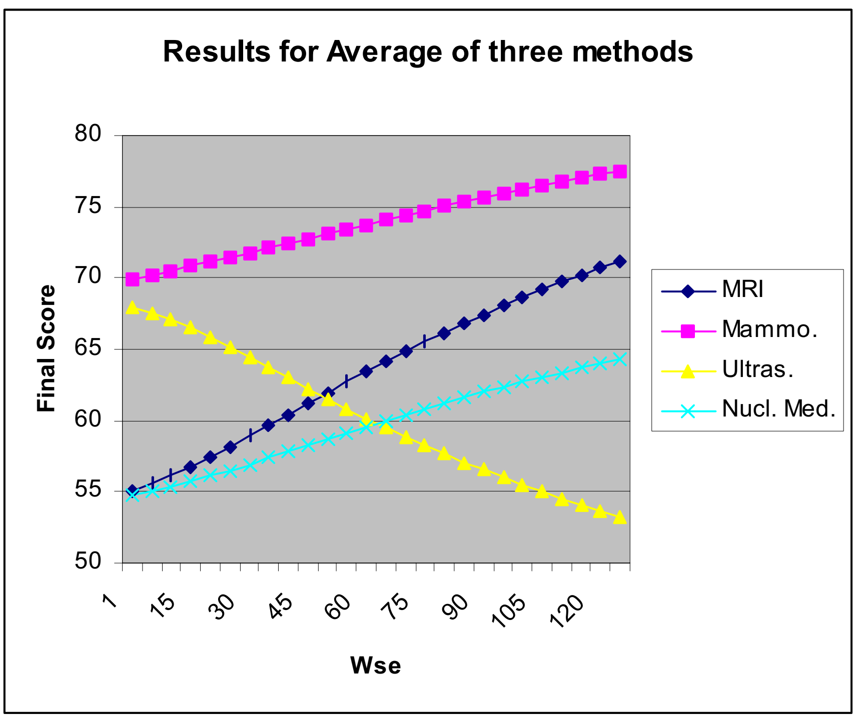
APPENDIX – CHARTS AND RESULTS

A1) *Effect of w_{SE} on Final Scores*

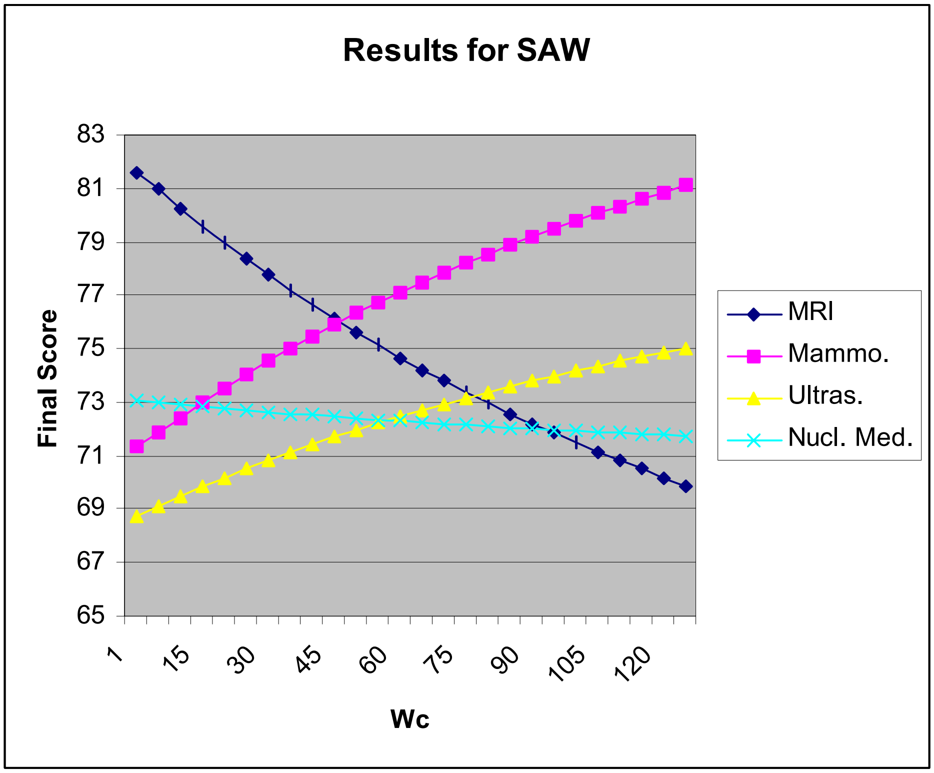


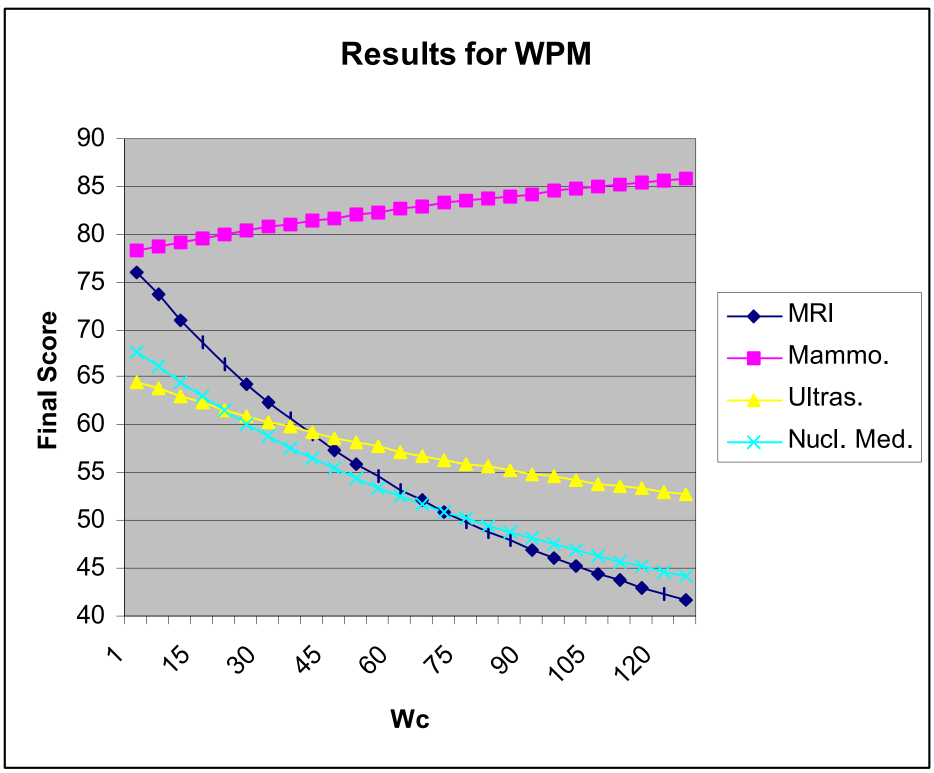


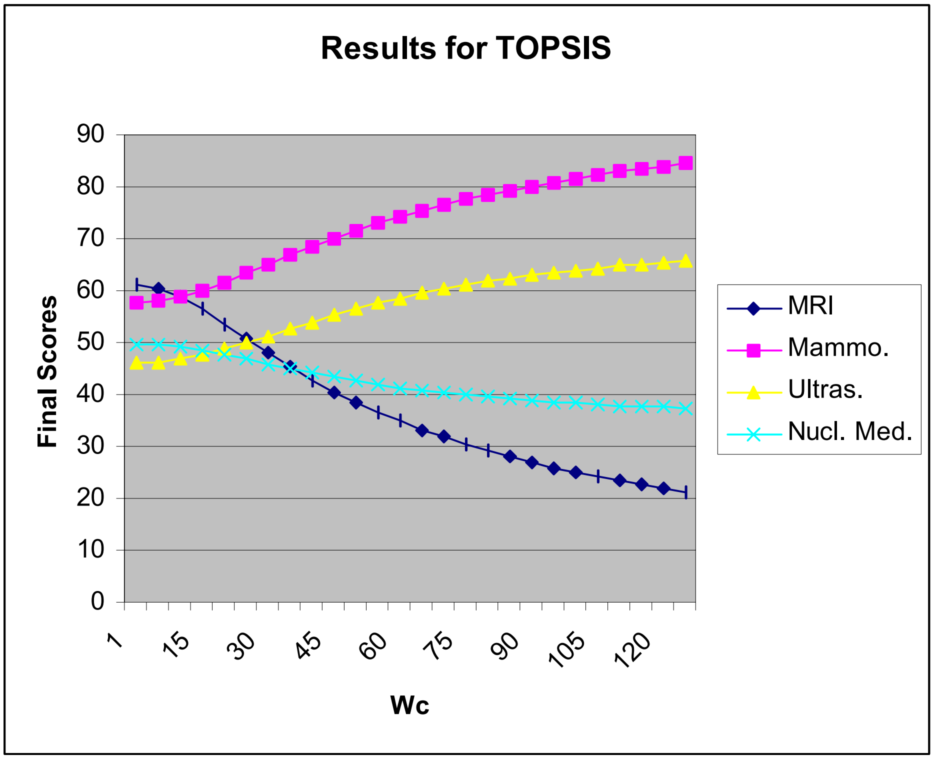

Results for TOPSIS
75
70
65
60
55
50
45
40
35
30
Final Score
1
15
30
45
60
75
90
105
120
Wse
MRI
Mammo.
Ultras.
Nucl. Med.

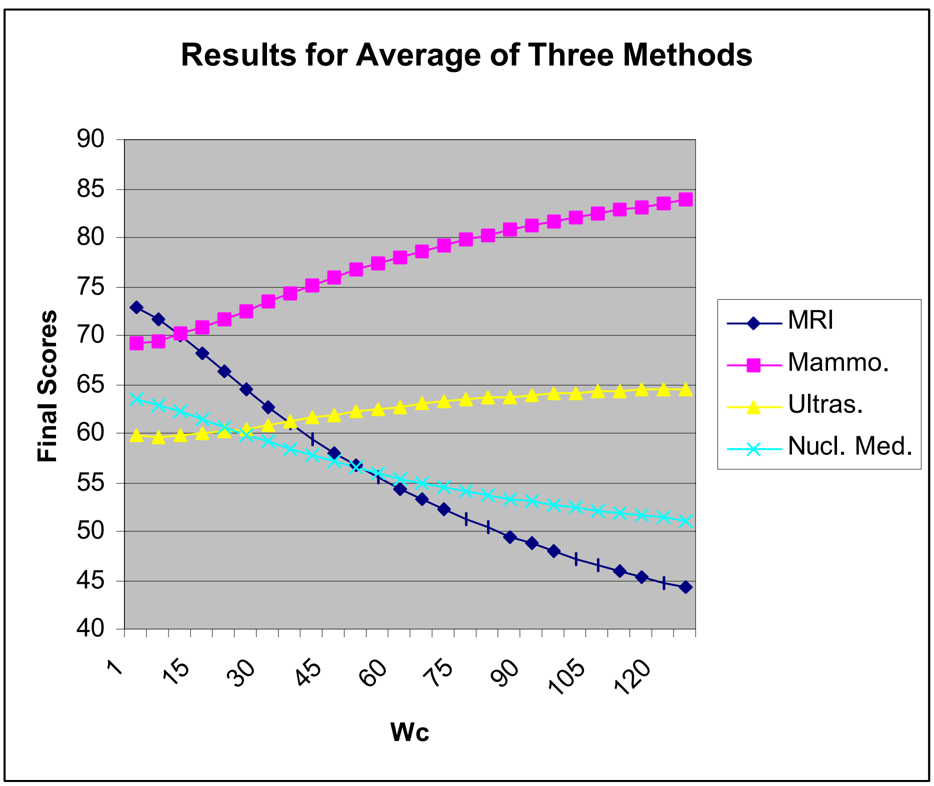

Results for Average of three methods
80
75
70
65
60
55
50
Final Score
1
15
30
45
60
75
90
105
120
Wse
MRI
Mammo.
Ultras.
Nucl. Med.

A2) ***Effect of w_C on Final Scores***

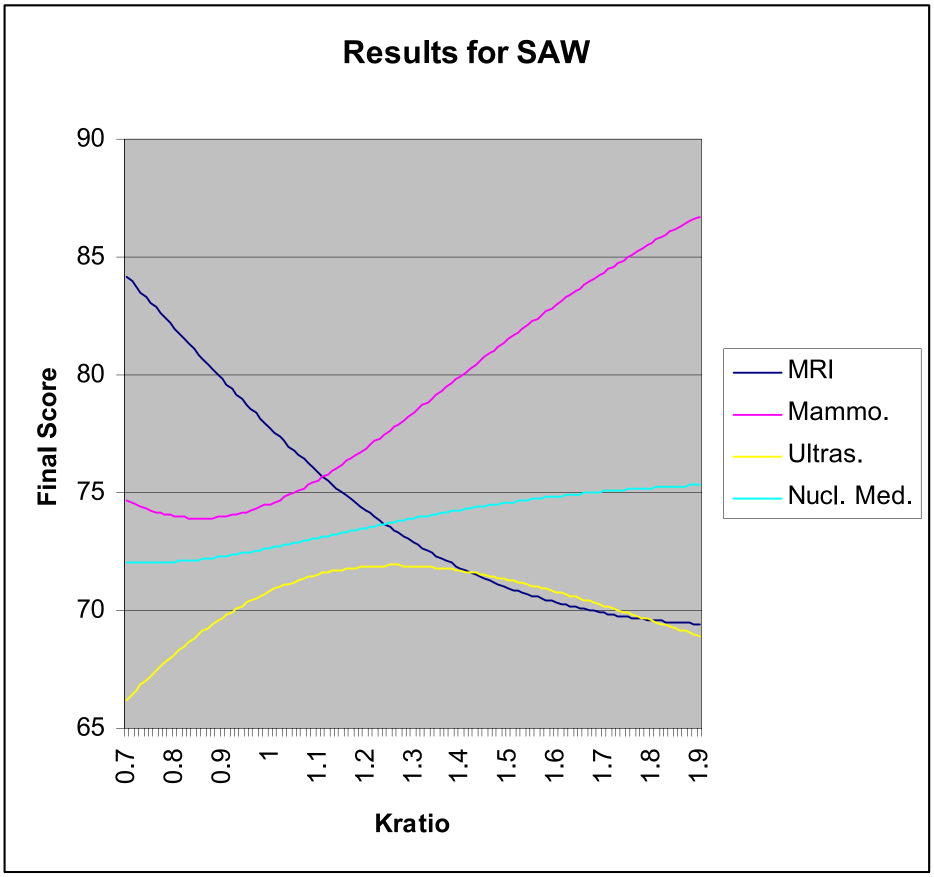


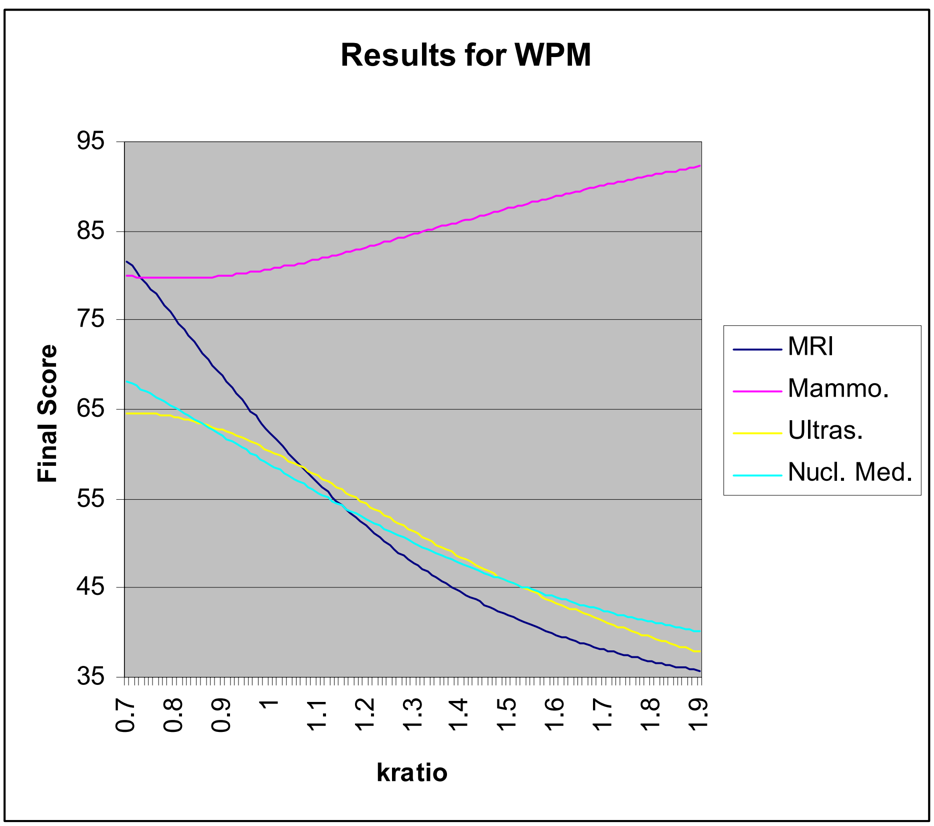


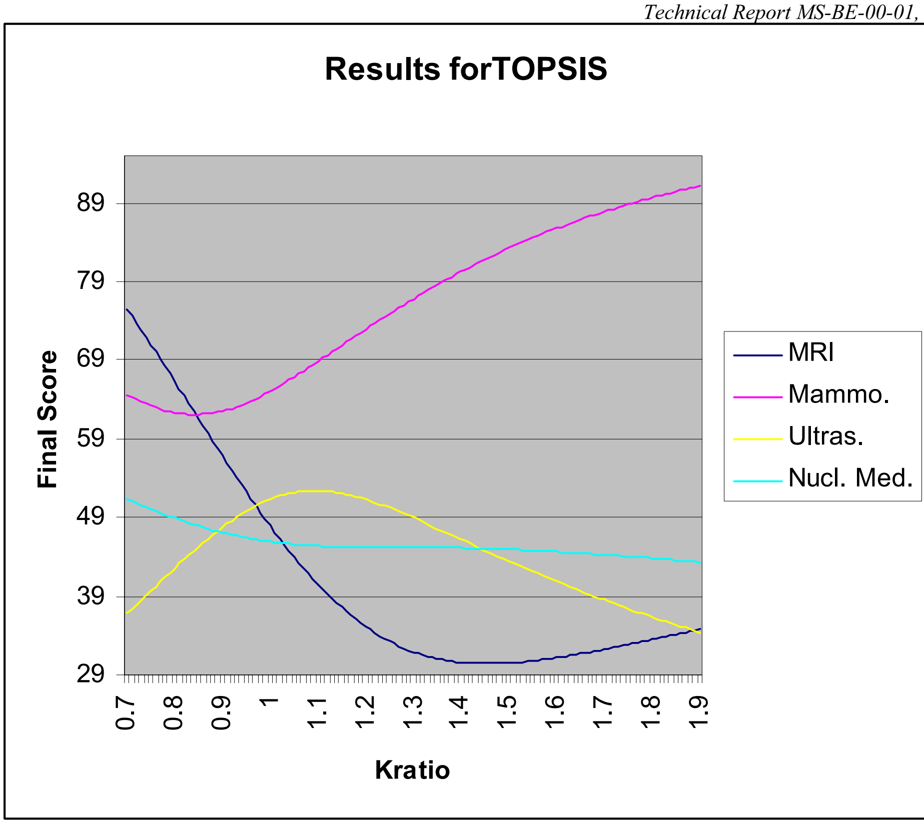

Results for TOPSIS
Final Scores
90
80
70
60
50
40
30
20
10
0
1
15
30
45
60
75
90
105
120
Wc
MRI
Mammo.
Ultras.
Nucl. Med.

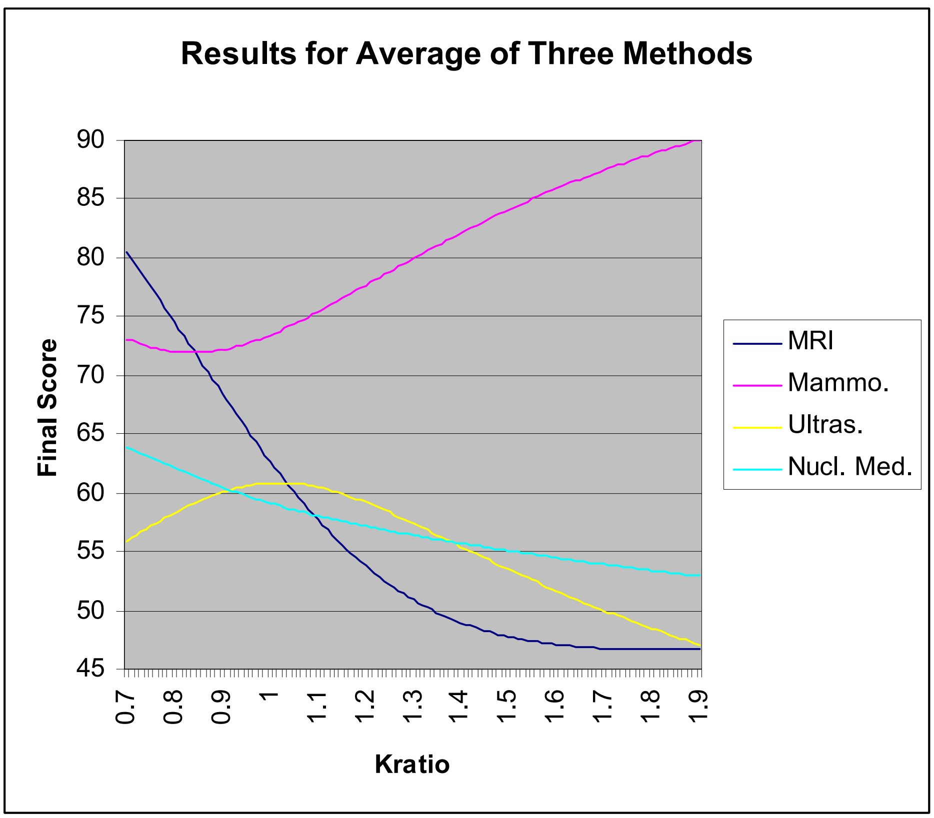

Results for Average of Three Methods
Final Scores
90
85
80
75
70
65
60
55
50
45
40
1
15
30
45
60
75
90
105
120
Wc
MRI
Mammo.
Ultras.
Nucl. Med.

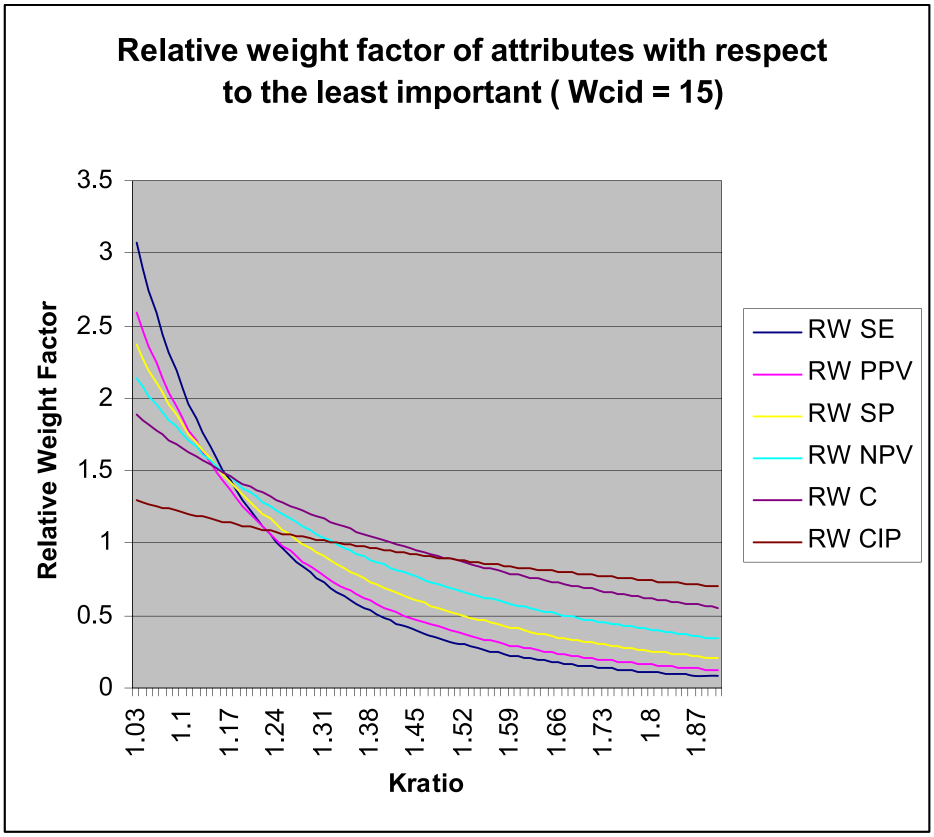
A3) ***Effect of Kratio on Scoring Methods, and on Relative Weight Factors***






Results forTOPSIS
Final Score
89
79
69
59
49
39
29
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
Kratio
MRI
Mammo.
Ultras.
Nucl. Med.


Results for Average of Three Methods
Final Score
90
85
80
75
70
65
60
55
50
45
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
Kratio
MRI
Mammo.
Ultras.
Nucl. Med.


Relative weight factor of attributes with respect to the least important (Wcid = 15)
Relative Weight Factor
3.5
3
2.5
2
1.5
1
0.5
0
1.03
1.1
1.17
1.24
1.31
1.38
1.45
1.52
1.59
1.66
1.73
1.8
1.87
Kratio
RW SE
RW PPV
RW SP
RW NPV
RW C
RW CIP

References:

[1] Kimbrough S., "Management Information Systems Notes, Decision Support", pp. 37-106, 1998

[2] Yoon K. P., Hwang C. L., "Multiple Attribute Decision Making, An Introduction", Sage University Papers (Series: Quantitative Applications in the Social Sciences), 1995

[3] Nunes L. W., Schnall M. D., Orel S. G., Hochman M. G., Langlotz C. P., Reynolds C. A., Torosian M. H., "Breast MR Imaging: Interpretation Model", Radiology Vol.202 pp.833-841, 1997

[4] Williams M. B., Pisano E. D., Schnall M. D., Fajardo L. L., "Future Directions in Imaging of Breast Diseases", Radiology Vol.206 pp.297-300, 1998

[5] Wu Y., Giger M. L., Doi K., Vyborny C. J., Schmidt R. A., Metz C. E., "Artificial Neural Networks in Mammography: Application to Decision Making in the Diagnosis of Breast Cancer", Radiology Vol. 187 pp.81-87, 1993

[6] Swets J. A., Getty D. J., Pickett R. M., D'Orsi C. J., Seltzer S. E., McNeil B. J., "Enhancing and Evaluating Diagnostic Accuracy", Medical Decision Making Vol.11 pp.9-18, 1991

[7] Khalkhali I., Mishkin F. S., Diggles L. E., Klein S. R., "Radionuclide-Guided Stereotactic Prebiopsy Localization of Nonpalpable Breast Lesions with Normal Mammograms", Journal of Nuclear Medicine Vol.38 pp.1019-1022, 1997

[8] Stavros A. T., Thickman D., Rapp C. L., Dennis M. A., Parker S. H., Sisney G.A., "Solid Breast Nodules: use of Sonography to Distinguish between Benign and Malignant Lesions", Radiology Vol.196 pp.123-134, 1995

[9] Hwang C. L., Yoon K., "Multiple Attribute Decision Making: Methods and Applications", Berlin/Heidelberg/New-York: Springer Verlag, 1981

[10] MacCrimmon K. R., "An Overview of Multiple Objective Decision Making", In J. L. Cochrane and M. Zeleny (Eds.), Multiple Criteria Decision Making (pp. 18-43), Columbia, SC: University of South Carolina Press, 1973

[11] Palmedo H., Biersack H. J., Lastoria S., et al., "Scintimammography with Tc-99m MIBI for breast Cancer Detection: Results of the Prospective European Multicenter Trial (abstr.), Journal of Nuclear Medicine Vol.38 pp.20P-21P, 1997

[12] Scopinaro F., Schillaci O., Ussof W., et al., "A Three Center Study on the Diagnostic Accuracy of 99mTc-MIBI Scintimammography", Anticancer Research Vol.17 pp.1631-1634, 1997

[13] Heywang S. H., Wolf A., Pruss E., et al., "MR Imaging of the Breast with Gd-DTPA: Use and Limitations", Radiology Vol.171 pp.95-103, 1989

[14] Kopans D. B., "The Positive Predictive value of Mammography", AJR Vol.158 pp.521-526, 1992

[15] Gisvold J.J., Martin J.K.Jr., "Prebiopsy Localization of nonpalpable Breast Lesions", AJR Vol.143 pp.477-481, 1984

[16] Rosenberg A.L., Schwartz G.F., Feig S.A., Patchefsky A.S., "Clinically Occult Breast Lesions: Localization and Significance", Radiology Vol.162 pp.167-170, 1987

[17] Bassett L.W., Liu T.H., Giuliano A.I., Gold R.H., "The Prevalence of Carcinoma in Palpable vs Impalpable, Mammographically Detected Lesions" (comment), AJR Vol.158 pp.688-689, 1992

[18] Khalkhali I., Villanueva-Meyer J., Edell S.L., et al., "Diagnostic Accuracy of Tc-99m-sestamibi Breast Imaging in Breast Cancer Detection" (abstr), Journal of Nuclear Medicine Vol.37 pp.74P, 1996

[19] Edwards W., "Use of Multiattribute Utility Measurement for Social Decision Making", in David E. Bell, Ralph L. Keeney, and Howard Raiffa, editors, *Conflicting Objectives in Decisions* pp.247-276, J. Wiley&Sons, New York, NY, 1977

[20] McIver J. P., Carmines E. G., "Unidimensional Scaling", Sage University Paper Series in Quantitative Applications in the Social Sciences pp.7-24, Newbury Park CA, 1981

[21] Spector P. E., “Summated Rating Scale Construction: An Introduction”, Sage University Paper Series on Quantitative Applications in the social sciences pp.7-82, Thousand Oaks CA, 1992

[22] Bridgeman P. W., “Dimensionless Analysis”, New Haven CT: Yale University Press, 1922