

# **Investigating the introduction of Spectroscopy and Diffusion-Weighted imaging (DWI) sequences to increase the specificity of Breast Magnetic Resonance Imaging (MRI):**

---

Thesis submitted in partial fulfillment of the requirements  
for the degree of Master of health science

**Peter Oberlin-Brown**

**1062833**

**18/07/2013**

**Unitec Institute of Technology**

**New Zealand**

## Declaration

**Name of candidate: Peter Oberlin-Brown**

This Thesis entitled: **Investigating the introduction of Spectroscopy and Diffusion-Weighted imaging (DWI) sequences to increase the specificity of Breast Magnetic Resonance Imaging (MRI)** is submitted in partial fulfilment for the requirements for the Unitec degree of master of health science.

### CANDIDATE'S DECLARATION

I confirm that:

- This Thesis/Dissertation/Research Project represents my own work;
- All information contained in this thesis by way of retrospective audit complies with the Unitec research and ethics regulations and was undertaken with institutional consent.

Candidate Signature: .....Date: **18/07/2013**

Student number: 1062833

## Abstract

Current breast MRI practice in New Zealand closely emulates the American and European guidelines. The most commonly used protocols used for imaging breasts using these guidelines are dynamic contrast enhanced (DCE) sequences that involve the intravenous injection of contrast and generation of contrast kinetic graphs to differentiate any suspicious lesions. Breast MRI has been widely researched and is described as having high sensitivity to breast lesions. Specificity of breast MRI (DCE) is more debatable with wider variations reported resulting in false positives.

This research has set out to examine the introduction of DWI and spectroscopy to routine protocols in order to evaluate their effect on sensitivity and specificity. This research was undertaken on a 1.5 Tesla MRI scanner using only pathology confirmed lesions. This selection of lesions included a wide range of all of the commonly presenting malignant lesions (Lobular, Invasive ductal, DCIS and Mucinous).

In order to fully evaluate the performance of DWI and spectroscopy an audit of 68 pathology confirmed breast lesions was undertaken. DCE MRI gave sensitivity and specificity of 90% and 62.5% respectively. When evaluating the same sample with DWI, specificity was increased to 88%. Perhaps the most significant finding of this research was 100% sensitivity to DCIS (n=11) when using DWI alone. Spectroscopy was undertaken using a smaller sample of mass-like lesions (n=21) achieving sensitivity of 100%. Positive predictive value for this sample was 90%.

In drawing conclusions from this research DWI using a b-value pairing of 0 and 750 mm<sup>2</sup>/s increases the specificity of breast MRI to malignancy in a clinically manageable time of less than four minutes and as such would be a worthy addition to routine breast MRI protocols. Spectroscopy demonstrated both high sensitivity and PPV to breast lesions but with a specificity of 66% and imaging times of between five and eight minutes needs further evaluation before being undertaken routinely.

## Acknowledgements

Breast cancer is a tragic disease which affects a significant proportion of the population if not directly then by association to friends and family.

This thesis is the culmination of two years' work and involves a topic which I believe can make a difference to the diagnosis, clinical management and follow up of women with breast cancer.

This work would not have been impossible without the assistance of my colleagues, tutors and family whose enthusiasm and encouragement have managed to keep me, focused on what has been a long journey. Thank you to Gillian, Adrienne and Suzanne for your guidance and assistance with the structure and technical aspects of my Master's thesis.

Thank you to Jill Oliver and Terrence Doyle for all your encouragement and expertise which allowed me to confidently settle upon this topic knowing that support would be at hand if required.

## List of Abbreviations

Abbreviation	Expansion/Definition
ACR	American College of Radiologists
ACS	American Cancer Society
ADC	Attenuation diffusion coefficient
ADH	Atypical ductal hyperplasia
b-Value	DWI weighting factor
BI-RADS	Breast Imaging-Reporting And Data System
BRCA	Breast cancer Gene mutation type 1 or 2
BREASE®	BREAst Spectroscopy Exam (GE healthcare)
CAD	Computer aided diagnosis
CADstream™	Commercially available CAD (Confirma) GE Healthcare
CSI	Chemical shift imaging
COMICE	Comparative Effectiveness of MR Imaging in Breast Cancer
DCE	Dynamic contrast enhancement
DCIS	Ductal carcinoma in-situ
DWI	Diffusion weighted imaging
EGFR	Epidermal growth factor receptor
EPI	Echo planar imaging
<sup>1</sup> H proton spectroscopy	Single voxel proton spectroscopy
HER2	Human epidermal growth factor receptor 2
IV	Intra-venous
ROI	Region of interest
LCIS	Lobular carcinoma in-situ
MRI	Magnetic resonance imaging
NAC	Neo-adjunct chemotherapy
NEX	Number of excitations (signal averages)

NMR	Nuclear magnetic resonance
NZHTA	New Zealand Health Technology Assessment
PET	Positron emission tomography
PRESS	Point resolved spectroscopy sequence
PPV	Positive predictive value
ppm	parts per million
SNR	Signal to noise ratio
STEAM	Stimulated echo acquisition mode
STIR	Short tau inversion recovery
tChol	Total choline peak
T2*	Gradient echo T2 weighted MRI sequence
TE	Time to echo
TR	Repetition time
VIBRANT™	Volume Imaging for BReast Assessment (GE healthcare)

## Table of Contents

Title .....	i
Declaration .....	ii
Abstract .....	iii
Acknowledgements .....	iv
List of abbreviations .....	v
Table of contents .....	vii
Table of figures, tables, graphs and examples .....	ix
<b>Chapter 1 Introduction.....</b>	<b>1</b>
<b>Chapter 2 Background .....</b>	<b>2</b>
DWI.....	2
Spectroscopy.....	4
<b>Chapter 3 Methodology .....</b>	<b>7</b>
Program evaluation .....	7
Research validity .....	8
<b>Chapter 4 Literature review .....</b>	<b>10</b>
MRI techniques .....	10
DWI.....	16
Spectroscopy .....	30
Summary.....	39
<b>Chapter 5 Study methods.....</b>	<b>42</b>
<b>Chapter 6 Results.....</b>	<b>47</b>
Examples .....	53
<b>Chapter 7 Discussion .....</b>	<b>61</b>
Current practice .....	61
DWI.....	67
Spectroscopy .....	72
<b>Chapter 8 Limitations.....</b>	<b>74</b>

<b>Chapter 9 Conclusion .....</b>	<b>75</b>
<b>Chapter 10 Recommendations.....</b>	<b>77</b>
<b>Chapter 11 References.....</b>	<b>78</b>
<b>Chapter 12 Appendix.....</b>	<b>83</b>



## Figures, Tables, Graphs and Examples.

### Figures

1	DWI + ADC demonstrating ROI placement.....	2
2	Spectroscopy voxel placement + Spectra.....	5
3	Mammogram showing micro-calcifications.....	13
4	CADstream™ subtraction/threshold with volume MIP (DCIS).....	14
5	CADstream™ subtraction/threshold with volume MIP (DCIS).....	15
6	ROC curve example.....	46

### Tables

1	Sensitivity and specificity Mammography + DCE MRI.....	11
2	Sensitivity and specificity Mammography, DCE, DWI + DCE & DWI.....	19
3	Sensitivity and specificity Mammography, DCE, DWI + DCE & DWI + Spectroscopy.....	32
4	Audit data (Pathology, DCE & DWI) .....	49
5	Audit data (Pathology, spectroscopy) .....	50

### Graphs

1	DWI ROC.....	51
2	ADC threshold comparison.....	52

### Examples

1	Chronic inflammatory change. False positive DWI result.....	53
2	Benign fibrotic change. False positive DCE & Spectroscopy result.....	53
3	DCIS. False negative DCE result.....	54
4	Lobular carcinoma.....	55
5	IDC. Positive DWI & Spectroscopy.....	56
6	IDC. Positive lesion result with DWI highlighting positive internal mammary lymph node.....	56
7	IDC. Positive lesion result.....	57
8	IDC. Positive lesion result.....	59
9	Papilloma. Benign Spectroscopy spectrum.....	60

## Introduction

Breast cancer is a highly emotive and tragic disease and it is the most common cancer affecting women in New Zealand with 'more than 2500 women diagnosed with the disease annually and the risk of a woman developing breast cancer at some time during her life being 1 in 9' (New Zealand Breast Cancer Foundation, 2010). Breast cancer screening programs in New Zealand predominantly use a mix of mammography and ultrasound. Breast magnetic resonance imaging MRI is a more recent addition to some imaging programs and where available is used at the discretion of that service. Typically those accessing breast MRI are women with genetic risk factors (BRCA-1 and -2 Gene mutations) and for the pre-surgical staging of those patients at risk of bilateral disease such as invasive breast cancers.

MRI breast imaging is used as an imaging tool because of its high sensitivity for detecting breast cancer (88 to 100%) (Yabuuchi, et al. 2008) demonstrating lesion morphology and contrast kinetics without the use of ionising radiation. However, while MRI demonstrates high sensitivity to breast cancer the majority of the literature available describes its specificity as being much lower (68 to 96%) (Yabuuchi, et al. 2008). As such, it has demonstrated potential weaknesses with false positive results misclassifying benign lesions as malignant. The consequences of any false positive diagnosis could include a patient undergoing unnecessary biopsy and/or treatment. As well as affecting diagnostic accuracy, false positive outcomes can cause additional emotional and physical stress for patients.

MRI is a rapidly developing imaging modality where new sequences and applications are constantly evolving and as such it is important for imaging protocols to change where improvements are evidence based. MRI protocols used in the imaging of suspected or confirmed breast cancer typically include the use of dynamic contrast enhanced (DCE) sequences. In addition, diffusion weighted imaging (DWI) and spectroscopy are two MRI sequences that have been widely researched and recently included into routine breast MRI protocols at the MRI centre where this research is to be carried out. As the introduction of DWI and spectroscopy sequences is a recent initiative it is the intention of this research to evaluate possible clinical impact these may have had. Should DWI and spectroscopy sequences be proven beneficial to breast MRI sensitivity and specificity, the findings of this project might then be used to make protocol recommendations whose validity could be relevant for other breast screening programs.

## Background

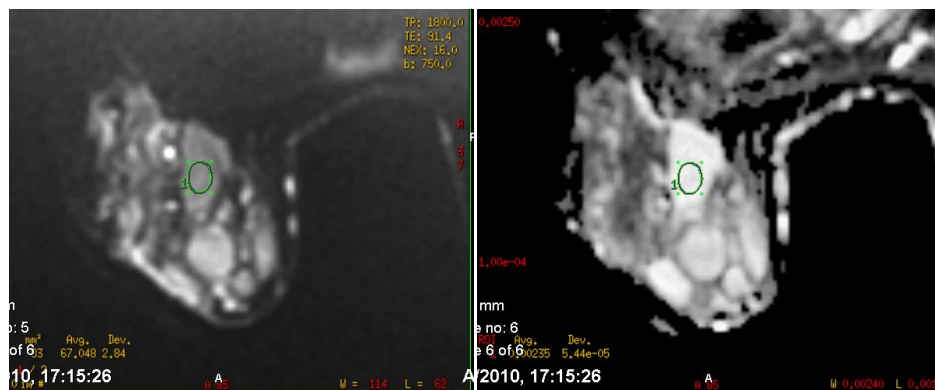
### DWI

DWI highlights lesion morphology and the key concept in this study is that it also allows quantitative characterisation of tumours by measuring apparent diffusion coefficient (ADC) values. ADC values can therefore be used to determine a lesion's classification (malignant or benign) using predetermined ADC threshold values.

DWI differs from routine breast MR techniques by making it possible to examine tissue at a microscopic or cellular level by measuring the mobility of water molecules within the cells.

The principle by which water molecules diffuse or move within cells in a given tissue is called Brownian motion (Bushong, 2010). Peters, et al. (2010), describe how the cellularity of a lesion can be indirectly assessed with DWI. The rationale is that, in malignant lesions, the cells are more densely packed than in benign lesions and normal tissue resulting in less space for Brownian motion of the water molecules leading to lower ADC values in malignant lesions compared with those in benign lesions and normal tissue.

**Figure 1: DWI with corresponding ADC map**



DWI

ADC

This corresponding pair of images (DWI and ADC) demonstrates a region of interest (ROI) most likely to be that of a cyst.

Lower or restricted diffusion appears as areas of high signal intensity on DWI. While restricted DWI helps with lesion conspicuity it is the corresponding low signal or low ADC

that is the quantitative index of the diffusivity of water molecules (Woodhams, Kakita, Hata, Iwabuchi, Umeoka, Mountford & Hatabu, 2009).

When performing DWI sequences, image weighting is controlled by a factor known as the b-value. The selection of the b-value effectively sensitises the sequence to cellular water by applying a strong magnetic field gradient pulse. The selection of the b-value is an important feature in the design of a DWI sequence. Lower b-values allow higher lesion conspicuity but are more susceptible to artefacts, while selection of higher b-values decreases artefacts allowing more accurate ADC characterisation. Also, while the direction of gradient application will not influence the results as these water molecules move randomly and areas of restricted diffusion will be demonstrated as areas of high signal intensity, the benefit in applying more directions (a maximum of three for two-dimensional imaging) is increased SNR (signal to noise ratio). As with all MRI sequences all parameters need to be optimised in order to achieve a balance between optimal SNR and scanning time (shorter times can result in less movement artifact).

When performing DWI sequences a minimum of two b-values must be used, one of which must be 0 in order to produce an ADC map and obtain quantitative values. When calculating ADC values the b=0 scan is termed as the registration scan by which the average signal intensity of the sampled tissue is divided when applying the ADC formula,

$$ADC = -(1/b) \ln(S_{DWI}/S_0)$$

(Partridge, DeMartini, Kurland, Eby, White & Lehman, 2010).

Pereira et al. (2009) have investigated the use of multiple and pairs of b-values in seeking to determine which were best for assessing breast tissue. The conclusion of this work was that the ADC obtained using a b-value pairing of 0 and 750 s/mm<sup>2</sup> gives a statistically reliable result with sensitivity of 92.3% and a specificity of 96.2%. Based upon the results of this research it is this b-value pairing that I will be using with the intention of evaluating and comparing these sensitivity and specificity values clinically. There is a large amount of literature including Yabuuchi, et al (2008) that advises the use of at least three other b-values, using a mean ADC value. While it is possible to run these three DWI sequences it makes the examination more difficult clinically and for patients due to the extra time penalties involved with running extra sequences, and as Pereira. et al. (2009) have demonstrated specificity does not appear to change. Each DWI sequence that I will be using takes three minutes and fifty

seconds, and an additional four minutes as opposed to twelve, is of more use practically to the standard twenty minute imaging protocol.

The intention of this research is to examine the use of DWI not as a stand alone procedure but in order to complement DCE imaging which, as the literature demonstrates, has very good sensitivity. It is the hypothesis of this research that the specificity of the DWI sequence is the most important factor and as such it is the ADC quantification that should be optimised in order to enhance any diagnosis of breast cancer.

### *Spectroscopy*

Spectroscopy is an MRI sequence that may be used specifically for tumour classification by demonstrating the metabolites of a selected region of interest (ROI). Spectroscopy is in effect an MRI technique which seeks to replicate a physical biopsy in a non-invasive method by analysing the chemical components or metabolites of a selected tissue.

The principles of MRI Spectroscopy arise from a technique widely used in the scientific analysis of chemicals known as nuclear magnetic resonance (NMR) spectroscopy. The primary principle in NMR is that each metabolite or nuclide has its own unique radiofrequency. By applying a Fourier transform to a received MRI signal it is possible to generate a frequency spectrum (Bushong, 2003). The position of the metabolites within a selected tissue is controlled by the tissues atomic structure or electron binding.

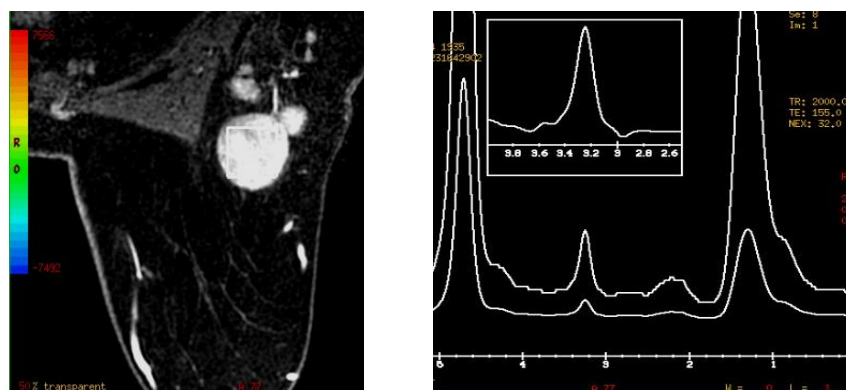
Breast MRI spectroscopy can be undertaken using single or multi-voxel sequences. As the descriptions suggest single voxel sequences interrogate a single voxel, while multi-voxel techniques involve a much larger sample area and require much less operator input (placement and lipid saturation) upon set up of the sequence. Multi-voxel techniques are more commonly undertaken at higher field strengths due to lipid contamination. At 1.5 Tesla multi-voxel techniques have been described as being difficult to diagnose (Su, 2008), while 3 Tesla scanners allow both better resolution and superior SNR with Gruber et al. (2011) reporting sensitivity and specificity of 97% and 84% respectively.

Clinically, single voxel spectroscopy techniques are the most common sequences undertaken; with the majority of the literature at 1.5 Tesla investigating these sequences qualitatively by visual identification of a choline peak. Quantitative techniques may be used by either SNR thresholds of absolute quantification which requires further spectroscopic techniques.

With breast MRI spectroscopy the presence of a choline metabolite peak is used to determine the pathological nature of the sample. When examining the rationale behind using the presence of a choline peak, Bartella et al, (2007) state that ‘the diagnostic value of proton (hydrogen 1 [ $^1\text{H}$ ]) MR spectroscopy is typically based on the detection of elevated levels of choline compounds, which are a marker of tumour’ (p. 80). This description of choline compounds refers to total or composite choline levels which are then assessed qualitatively on a yes/no basis.

Tozaki and Fukuma (2009) describe the various metabolites that make up these composite choline levels in more detail stating that they are best demonstrated using water as a reference (positioned at 4.7ppm). It is the presence of phosphocholine at 3.22-3.23ppm that defines malignancy while other metabolites such as glycerophosphocholine, taurine, and myo-inositol were positioned between 3.27-3.28ppm and deemed to be benign.

**Figure 2: Spectroscopy voxel and corresponding Spectrum**



Spectroscopy voxel placement      Spectroscopy (Phosphocholine peak on inset)

There is a large amount of variation regarding the sensitivity and specificity in the literature of Proton ( $^1\text{H}$ ) MR spectroscopy (MRS). A recent study by Tozaki and Fukuma (2009) has demonstrated that Proton ( $^1\text{H}$ ) MRI spectroscopy was applied for mass lesions larger than 15 mm, with a diagnostic sensitivity and specificity of 82% (28/34) and 69% (11/16), respectively, while Bolan, Nelson, Yee and Garwood (2005) demonstrated an increase in specificity from 62.5% to 87.5%. There is debate as to the usefulness of  $^1\text{H}$  MRS in assessing contrast enhancing non-mass lesions with Bartella et al.(2007) finding 100% sensitivity and 85% specificity. Tozaki and Fukuma state that the sample used by Bartella did not contain

enough patients with DCIS and contradict these findings in a similar study with 57% of the sample containing patients with DCIS which demonstrated a false negative rate of 68% when using MRI spectroscopy on contrast enhancing non-mass lesions. The significance of this literature is that in this study I will be only assessing spectroscopy sequences undertaken to evaluate mass lesions larger than 15mm.

The MRI spectroscopy sequences used in this research varied in time depending upon the size of the tissue sampled. Tissue samples in spectroscopy are defined by voxels, with a typical spectroscopy voxel being a three dimensional cube. A 20mm<sup>3</sup> voxel takes approximately five minutes to acquire, while 18mm<sup>3</sup> takes eight minutes and 15mm<sup>3</sup> takes twelve minutes. Due to the significant time variations the largest possible voxel size is typically used for each lesion sampled. Penalties for using voxels that are too large include lipid contamination or side banding of the spectrum. The spectroscopy sequences used in this research, while part of routine protocols, were performed on a case by case basis determined by the lesion size, imaging time and patient comfort considerations.

## Methodology:

### *Program evaluation*

In order to evaluate existing breast MRI protocols this research was undertaken using the structured approach of the programme evaluation research method. Using the quantitative assessment of sensitivity and specificity along with the qualitative assessment of the associated literature this research methodology allows recommendations to be made about whether these protocols remain the same or identified whether changes are appropriate.

Programme evaluation is a specific form of the research method more commonly known as evaluation research. It is used as a tool in evaluating current and prospective procedures and, by the use of either quantitative data or qualitative interpretation; it seeks to make recommendations for improvement. Hart (2006) defines evaluation research methods as being a “specific assessment for the purposes of making recommendations for change to a policy, programme or product using whatever data or collection tools and approach is appropriate” (p.330).

Most of the literature surrounding programme evaluation research methods appears to primarily be for educational purposes, but upon closer examination of this structured approach its validity appears relevant to evaluating any specific protocol or program. It is this structured approach which would simplify procedures for examining my hypothesis, allowing me to assess the existing breast imaging programme rationale and examine the effects of diffusion and spectroscopy sequences. The flexibility of programme evaluation also means that this hypothesis could be examined by the use of either quantitative or qualitative data methods before making recommendations for change.

In order to best appreciate the structure of a typical programme evaluation research study I have chosen to simplify this into steps:

- Formulation of an initial criteria or hypothesis for the testing of an existing or proposed program.
- Data sampling undertaken using a variety of qualitative and/or quantitative methods as long as the final outcome involves making recommendations.
- Data collection and analysis.
- Structured conclusion with clear recommendations.



Program evaluation research methods are typically either summative or formative. It is the summative type of evaluation research that appears to be the best fit for assessing existing programme criteria such as that which I am proposing, with Hart stating that it “aims to evaluate ‘effectiveness and need’ so as to provide data for decisions on the continuance, changes to, or merger of a policy or programme” (2006, p. 330).

Trochim (2006) discusses evaluating between observational and correlational methods and the effectiveness of quasi-experimental and experimental designs for determining whether observed effects can reasonably be attributed to the intervention and not to other contributing sources. It is in this manner described by Trochim (2006) that I would wish to examine my hypothesis and with the incorporation of both observational and correlational methods am seeking to provide evidence for any potential protocol recommendations.

### *Research validity*

A limitation or weakness that programme evaluation research methods appear to have is a perceived lack of external validity. This validity issue relates directly to the fact that this method typically involves examination of a specific programme and commonly involves evaluation of this programme in its context, leading to questions about whether the findings are applicable to other applications. Anderson states that these perceived validity issues have caused people to incorrectly “raise the question of whether evaluation is indeed research” (1998, p. 136).

Trochim defines validity as referring “to the approximate truth of conclusions” further stating that “external validity is the degree to which the conclusions in your study would hold for other persons in other places and at other times” (2006).

Peat (2001) examined research validity from a health science perspective describing how validity is often confused with repeatability and how external validity can be measured by means of sensitivity analysis and subjective judgement. Internal validity is a more complex topic again and, where measurement is required, involves statistical analysis of the research criterion, construct, face and content in order to for comparisons to be made.

From the previous definitions of validity when applying these to a research situation, internal validity is crucial to a research project ensuring that ‘there are no errors internal to the design of the research method’ (Neuman, 2003, p. 187), while external validity is dependent on the intended application of the research.

The issue of external validity is a difficult one for my hypothesis. The focus or context of this study is intended to be for a single centre only and as such, threaten its external validity both locally and nationally. While external validity is a difficult issue for this particular research, should this research show promise in improving breast MRI specificity then it could potentially be used as a model for similar studies at other centres.

Another potential weakness in providing specific recommendations in the programme evaluation method is that where existing protocols are proven to be sufficient, any recommendation of continuing with the status quo would appear to be a failure on the part of the research given the methods focus upon change. This focus upon the need to change is not necessarily a failure if the research shows there is no need for change.

Cohen (2000) supports this perceived weakness by stating that:

“most evaluation studies in which tests of statistical significance show no support for the study’s hypothesis are never published” further stating that “over time this publication bias tends to skew the education literature toward an overestimate of the effectiveness of particular programs” (p. 455).

The issues in not providing recommendations for change are also a concern in utilising the program research format for my hypothesis. Any failure to make recommendations for change could be conceived in many ways. One perception would be that there would be little benefit in submitting research which didn’t successfully achieve the structure it was seeking to. Failure to make recommendations for change can also be perceived as support for the status quo. While support for the status quo may be perceived by some as a weakness this support may also answer the question. A third perception of not providing recommendations for change when using this methodology could be where further research is required before any decisions may be made.

While there are stated weaknesses in programme evaluation methods, once identified I believe that these can be overcome by placing further emphasis on a project’s design in order to prevent the research from failing or having to be manipulated in a manner which would affect both the research validity and credibility.

## Literature review

It is the purpose of this literature review to examine the current state of breast MRI and examine the literature relating to the clinical performance of breast DWI and Spectroscopy.

### *MRI Techniques*

Currently, the most commonly used breast MRI sequences for the evaluation of breast cancer are those using dynamic contrast enhancement (DCE) (Beatty, 2007). DCE techniques require intravenous (IV) contrast administration and a fat saturated contrast sensitive sequence (T1 weighted) with multiple phases imaged at predetermined time intervals. These multiple phases' allow post processing and generation of contrast kinetic or time intensity graphs that can then be used to characterise a lesion as being either benign or malignant.

Contrast kinetics can be analysed using generic MRI scanner or second party Computer Aided Diagnosis (CAD) software, which allows colour overlays, vascular mapping, volumetric and positional data as well as automated image reformatting. CAD breast imaging analysis is a newer technique that is not available to all centres that undertake breast MRI.

The introduction of new post processing analysis techniques such as CAD and sequences such as DWI and/or spectroscopy can potentially affect the sensitivity and specificity of breast MRI and need consideration in evaluating the overall performance of this technique.

### *Sensitivity and specificity*

In evaluating breast MRI the standard manner by which techniques are compared is by direct comparison of sensitivity and specificity. The sensitivity of an examination is the probability that results at imaging are positive in those patients who have the disease, which in the context of breast MRI is the number of breast lesions detected expressed as percentage of the total number of lesions known to be present within the same sample (pathology). Specificity is the probability that results at imaging are negative in patients who do not have the disease (Huang et al. 2004).

**Table 1 - Comparative sensitivity and specificity breast imaging table (Mammography vs. Breast MRI).**

	No. Of Malignant lesions	No. Of Benign lesions	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Authors
<b>Mammo</b>	167 DCIS  89 high grade DCIS	n/a	56%  52%	n/a	55%	n/a	Kuhl (2007)
<b>MRI DCE</b>	167 DCIS  89 high grade DCIS	n/a	92%  98%	n/a	59%	n/a	Kuhl (2007)
	30	20	100%	62.5%	n/a	n/a	Huang et al. (2004)

Abbreviations: DCE = Dynamic contrast enhanced MRI, DCIS = Ductal carcinoma in-situ, NPV = Negative predictive value, PPV =Positive predictive value, Mammo = Mammography.

Standard MRI techniques using DCE to evaluate breast cancer demonstrates very high sensitivity to breast cancers (88 to 100%), with specificity being reported as potentially having significantly lower values (68 to 96%) (Yabuuchi et al. 2008). The values of sensitivity and specificity by Yabuuchi et al. are ranges quoted in a meta-analysis of previous research in order to give context and as such are useful here in describing the sensitivity and specificity of DCE breast MRI across a number of studies.

In original research that included the assessment of 50 breast lesions after positive findings at mammography, Huang et al (2004) quoted 100% sensitivity and 62.5% specificity (see table 1). This research involved a relatively small sample (n=50) and involved only lesions identified by positive mammography results. Because of this selection bias towards confirmed cancers this has affected the number of benign lesions available to calculate specificity.

Because it has a lower specificity than sensitivity, DCE breast MRI demonstrates weaknesses with false positive results, potentially misclassifying benign lesions as malignant. The consequences of any false positive diagnosis could include a patient undergoing unnecessary

biopsy and/or treatment. As well as affecting diagnostic accuracy, false positive outcomes can also cause additional emotional and physical stress for patients, some of whom have a strong familial history of breast cancer.

Williams et al (2007) have investigated the clinical use of Computer Aided Diagnosis (CAD) demonstrating both increases in the sensitivity and specificity of clinical breast MRI for evaluating benign from malignant lesions. These authors showed that CAD significantly reduced their false positive rate by 8.8%.

In a meta-analysis of CAD in breast MRI, Dorrius et al. (2011) manage to both agree and contradict with Williams et al. by stating that CAD in breast MRI has little influence (less than one percent) on the sensitivity and specificity of experienced radiologist, while at the same time finding an increase in sensitivity of 17% when the same sample was evaluated by radiology residents or registrars.

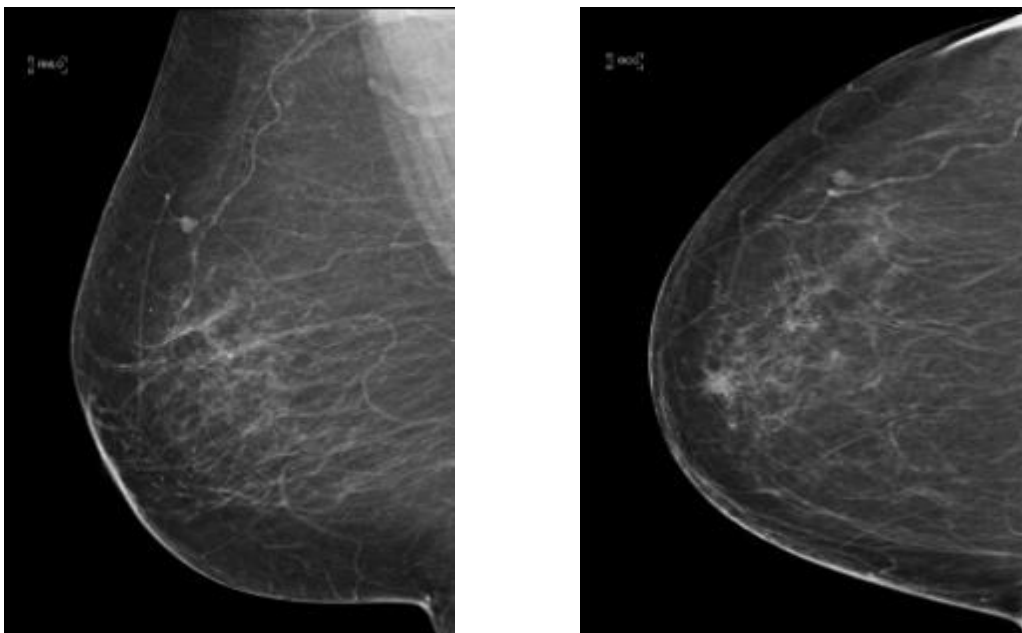
#### *Earlier Detection or Changing the Way Breast Cancer is Detected*

With increases in the sensitivity and specificity of breast MRI with new technologies such as CAD, breast MRI is now detecting lesions that in the past may have been overlooked. Historically, before mammography, only 3% of all cancers diagnosed were DCIS while with current breast screening practices the figure is now closer to 20% (Hall, 2010). Turnbull et al. (2010) questioned the relevance of identifying additional smaller foci stating that in the COMICE trial preoperative breast MRI made no difference to re-excision rates. In contrast, breast MRI in evaluating DCIS is advocated by Mossa-Bassa (2010) who, despite the COMICE trial, state that the early diagnosis of DCIS is important, describing how a large proportion of intra-ductal cancers can progress to invasive carcinoma.

DCIS can be present in different grades with high or intermediate grade DCIS able to easily spread invasively to other breast tissues before becoming metastatic. Low-grade DCIS and atypical ductal hyperplasia (ADH) are often considered to be insignificant as, when contained within the milk ducts of the breast, it does not have the ability to enter the patient's blood or lymphatic systems and thus does not become metastatic. The importance of identifying low-grade in-situ or indolent cancers such as low-grade DCIS and ADH is discussed by Hall (2010) who describe how recent genetic and molecular data has shown that they are more closely related to lobular carcinoma in-situ (LCIS) than to high-grade DCIS, confirming the docility described by Kuhl et al. (2007) (see table 1).

The identification of DCIS has typically been the domain of mammography due to the visualisation of calcium deposits, which occurs because DCIS obtains nutrients by diffusion fed from vessels outside the epithelial lining of the milk ducts that it is contained within. Hall (2010) describes that when the ducts become so tightly packed the diffusion distance becomes too large to sustain all of the cells causing hypoxia necrosis and ultimately microcalcifications within some of the lesion. DCIS calcification does occur in the majority but not all lesions and where there is no lesion calcification it becomes invisible to mammography.

**Figure 3: Mammography**



Right breast Mammogram demonstrating micro-calcifications

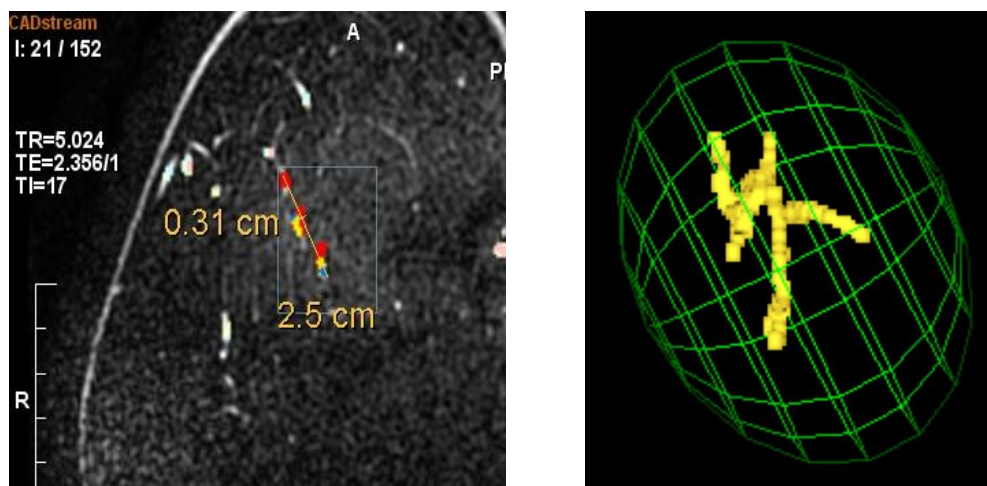
(Figures 3,4 & 5 relate to the same patient with micro-calcifications and DCIS)

In a study of 167 women diagnosed with pure DCIS, Kuhl et al. (2007) evaluated the sensitivity of DCE breast MRI versus the same cohort with mammography only. This study involved a retrospective audit and involved patient data collected over a five year period, with the authors seeking to investigate the sensitivity of each method of detection with regards to the biological profile of each lesion. The findings of this study were that 56% were detected by mammography, while 92% were seen using MRI alone. Of the 167 patients, 89 were pathology confirmed as being high-grade with the sensitivity of mammography decreasing with higher nuclear grading and 48% of these missed by mammography alone.

Specificity was not reported in this research and positive predictive value was reported to be 59% for MRI compared to 55% for mammography alone. If the findings of this study were repeatable (in different centres) the findings would be significant with MRI potentially preventing up to 50% of lesions that are likely to become invasive breast cancers.

In an audit of the New Zealand national breast screening register (ORION Healthcare), examined DCIS on breast MRI and showed that when assessing patients with intermediate and high-grade DCIS using MR that there is a 20% chance of detecting an invasive cancer that would otherwise have been missed using conventional mammography. Also identified in this audit was that high-grade DCIS presenting without calcification had an even higher chance (28%) of being associated with an invasive cancer.

**Figure 4: MRI CADstream**



MRI CADstream™ subtraction/threshold image with corresponding volume MIP of the same right breast depicted in the previous mammogram. Volume and threshold images demonstrate a ductal pattern consistent with the previous mammogram and pathology confirmation of DCIS.

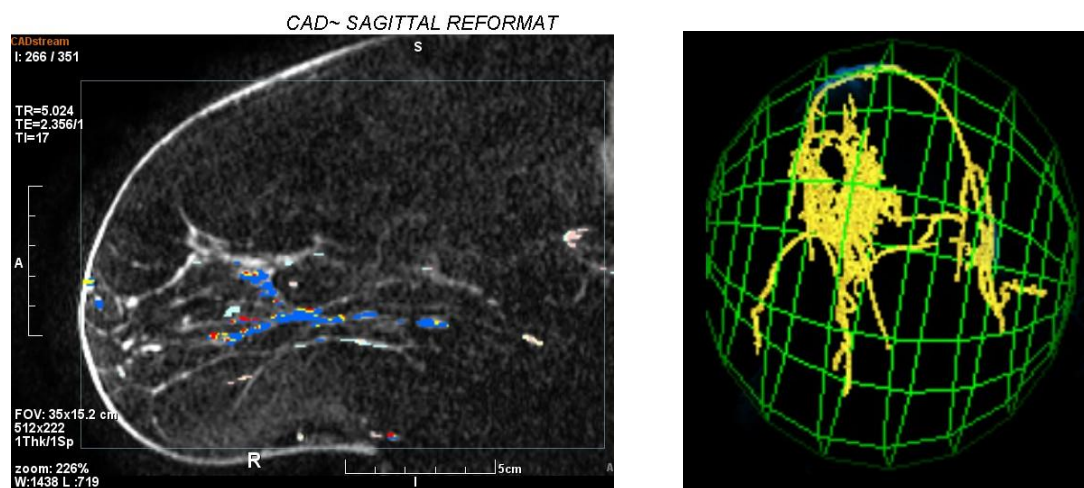
Boyd (2011) describes how that increased neovascularity, which invasive lesions exhibit, DCIS is associated with increased protease activity being fed by diffusion across the epithelial tissue of the ducts in which they are contained. Higher-grade DCIS lesions were described as having correspondingly higher levels of protease activity with Kuhl (2009) describing how gadolinium contrast media does not permeate into normal milk ducts but will perfuse into ducts containing DCIS. One of the difficulties in detecting DCIS with breast

MRI is that it is contained within the ducts and where there are no changes in morphology and CAD is not available these lesions can often be overlooked.

Conventional DCE breast MRI shows excellent sensitivity (96%) to breast cancer but is potentially problematic with some specificities being reported as being much lower (62.5%) (Huang et al. 2004) giving rise to false positives and additional biopsies. One way by which specificity has been reportedly increased (8.8%) is by using new techniques such as CAD post processing (Williams, 2007). CAD has also been reported to have significant effects on the sensitivity of DCE breast MRI for less experienced radiologists (Dorrius, 2011).

The identification of DCIS on conventional DCE MRI has been a contentious issue with some in the literature (Hall, 2010) stating that it is now being over diagnosed. MRI appears to be very good at identifying intermediate and high-grade DCIS which is relevant as it is in these forms that it can easily proceed to invasive disease. Some in the literature theorise that low-grade DCIS is more closely associated with lobular carcinoma in-situ (LCIS) than to intermediate and high-grade DCIS (Hall, 2010). As MRI has a much higher sensitivity of high and intermediate-grade DCIS should DCIS identified on MRI only be treated differently to that identified on mammography alone?

**Figure 5: MRI CADstream**



Further MRI CADstream™ threshold/volume reformats depicting ductal patterns of contrast uptake.



Diffusion-weighted imaging (DWI):

DWI of the breast is a newer MRI sequence that is yet to be introduced clinically at many centres. Since its emergence a significant amount of literature has been published evaluating both the technique as a standalone sequence and as an adjunct to conventional breast MRI (DCE).

Partridge et al. (2009) investigated the use of quantitative DWI as an adjunct to conventional breast MRI. This study involved a retrospective study of 70 women with 83 lesions demonstrated on DCE MRI who underwent biopsy. Positive predictive value (PPV) was then calculated for DCE-MRI alone and DCE-MRI plus DWI and the results compared by lesion type (mass and non-mass like enhancement).

The b-value pairing used by these authors were, 0 and  $600\text{s/mm}^2$  which was described as being the optimal b-value and giving the highest signal to noise ratio possible with a spin-echo diffusion-weighted sequence equal to  $1.1/\text{ADC}$ . This ADC value of 1.1 was selected by these authors following the review of an article by Jones (1999), whose article was based upon brain DWI. In searching the literature, this is the only study to compare SNR versus ADC ratios and while potentially relevant in that context, the benefits of using breast coils in this setting, and specifically imaging breast tissue, remains unproven, further to this, patients with tumours that were biopsy proven (B-RADS 6) before their MRI were excluded in this research (Partridge et al., 2009). The authors, while listing this exclusion, have not explained their reasons for doing so, even though the data used was assessed retrospectively and could have been controlled for bias. The authors stated that the purpose of this study was to improve the specificity or ability of breast MRI to correctly predict biopsy outcome and it was this outcome which probably influenced them into excluding these lesions. Inclusion of these biopsy proven lesions would have potentially increased the sensitivity of the examination and their inclusion wouldn't have made any clinical difference as the study involved a retrospective analysis.

Partridge et al. (2009) described PPV increases from 37% to 47% when DWI was included, but reported problems with an overlap of benign and malignant ADC thresholds affecting further increases. The authors used a slice thickness of 5mm in order to get enough SNR and suggested that this caused problems with partial voluming and difficulty in assessing some lesions. The use of a 5mm slice thickness appears to be standard practice and the problems described with threshold overlap were more than likely intrinsic problems associated with the

b-value selection of  $600\text{s/mm}^2$ . Pereira, et al. (2009) describe lower b-value selections and the associated ADC thresholds as being more likely to be influenced by perfusion effects, and overcome by the use of a larger b-value.

More recently, Partridge et al. investigated the effects of lesion type and size when discriminating benign and malignant lesions using ADC values (Partridge et al. 2010) in 91 women with 116 breast lesions identified by DCE breast MRI (see table 2). The authors again used b-values of 0 and  $600\text{ s/mm}^2$ . The primary intention was to determine whether lesion size or type (mass like or non-mass like) affected ADC values in the differentiation between malignant and benign lesions. The results of this study showed that there is no relationship between lesion size and ADC values, with the authors stating that their sample included 53/116 lesions less than one centimetre in size. The authors also demonstrated that DWI holds promise in assessing non-mass like lesions. In addition, the authors discussed their previous work in predicting PPV, stating that it was for use in implementing an ADC threshold into biopsy recommendation criteria for suspicious lesions and how this may improve the overall PPV of breast MRI.

Partridge (2009 and 2010) have evaluated breast DWI sequences using both PPV and specificity as the way in which they present their findings. Both articles were published in separate peer reviewed journals involving a retrospective audit of the same cohort of patients and as such are effectively the same research in different wrappers.

In a meta-analysis of twelve peer-reviewed articles on the use of ADC in the differential diagnosis of breast tumours, Tsushima, Takahashi-Taketomi, & Endo, K, (2009) have grouped together studies in table form allowing easy comparison of scanner type, b-values, ADC thresholds and other technical factors used in each study. Statistical analyses were performed in order to allow comparative sensitivity and specificity. The overall findings of this meta-analysis were a pooled sensitivity of 0.89 and specificity of 0.77 with the authors recommending a maximum b-value pairing of  $1000\text{ s/mm}^2$  using a threshold of  $1.23 \times 10^{-3}\text{ mm}^2/\text{s}$ . Of the twelve articles reviewed seven used a b-value pairing of  $1000\text{ s/mm}^2$ . While it is useful to see comparisons such as this in the literature it is difficult to know how to apply their conclusions as no two of the included studies use the same parameters and as such show large variations between each of the individual studies, which the authors themselves commented on describing this as being notable heterogeneity.

Similar research (Yabuuchi, et al., 2008) that was not included in the previously mentioned meta-analysis also examined the use of DWI sequences as an adjunct to DCE imaging. This research used a significantly larger patient cohort than that used by Partridge (2009 and 2010) initially examining 177 patients with 192 breast lesions. The authors assessed DWI in combination with DCE breast MRI commenting on sensitivity and specificity as well as PPV (see table 2). Combinations of three b-values were used, with these being 0, 500 and 1000 s/mm<sup>2</sup>. The results of this study showed that when using an ADC threshold of less than  $1.1 \times 10^{-3}$  that a sensitivity of 92%, specificity of 86% and PPV of 97% were achievable. The authors of this work stated that a large proportion of benign masses were excluded due to no pathology being available. This exclusion was described as a sample bias which would have potentially affected the specificity either directly or by influencing confidence levels. The specificity in this work was at the higher end of what is quoted in the literature and more than likely due to the predominance of invasive ductal carcinoma in the malignant lesions sampled (148/161).

In contrast to the high specificity of Yabuuchi, et al., (2008), Woodhams et al. (2005) quoted sensitivity of 96% and specificity of only 46% when evaluating 76 pathology confirmed breast lesions with quantitative DWI. This study used a DWI sequence with b-value pairings of 0 and 750 mm<sup>2</sup>/s. This lack of specificity is significantly less than reported by all of the articles previously reviewed in this study and is perhaps in part due to the ADC threshold of  $1.6 \times 10^{-3}$  mm<sup>2</sup>/s, which is significantly higher than the average threshold of  $1.23 \times 10^{-3}$  mm<sup>2</sup>/s quoted by Tsushima, Takahashi-Taketomi and Endo (2009) This high sensitivity and low specificity was also discussed by Marini et al. (2007) who found that more reliable specificity results were achieved by lowering ADC thresholds.

Woodhams et al. discuss the choice of b-values stating that malignant tumours exhibit greater perfusion effects than those that are benign due to angiogenesis and the proliferation of blood vessels surrounding these lesions, theorising that because of these effects there could be greater offset between the ADC thresholds of these two groups. Longer b-values require longer pulse sequences (TE's) which could be more susceptible to distortion or movement artefacts.

These authors also describe their theory that lesions containing or surrounded by hemosiderin and/or haemorrhage are more susceptible to magnetic susceptibility artefact, which can increase the ADC value and affect specificity. This magnetic susceptibility was blamed for a

false negative ADC reading within their patient cohort. Blood or hemosiderin deposition is generally well visualised on T1 weighted scans with this more than likely accounted for if a combined DCE and DWI analysis had taken place.

*Pathology specific sensitivity and specificity*

**Table 1: Comparative sensitivity and specificity breast imaging table (mammography vs. Breast MRI with and without DWI sequences).**

	No. Of Malignant lesions	No. Of Benign lesions	b-Value mm <sup>2</sup> /s	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Thres holds mm <sup>2</sup> /s	Authors
<b>Mammo</b>	167 DCIS  89 high grade DCIS	n/a	n/a	56%  52%	n/a	55%	n/a	n/a	Kuhl (2007)
<b>MRI DCE</b>	167 DCIS  89 high grade DCIS	n/a	n/a	92%  98%	n/a	59%	n/a	n/a	Kuhl (2007)
	30	20	n/a	100%	62.5%	n/a	n/a	n/a	Huang et al. (2004)
<b>DWI</b>	29	87	0 & 600	96%	56%	39%	98%	1.6 x 10 <sup>-3</sup>	Partridge et al. (2010)
<b>DWI&amp;DCE</b>	61	14	0, 500 & 1000	92%	86%	97%	71%	1.1 x 10 <sup>-3</sup>	Yabuuchi et al. (2008)
	26 (2 DCIS)	26	0 & 750	92.3%	96.2%	n/a	n/a	1.24 x 10 <sup>-3</sup>	Pereira et al. (2009)
	31 (1 DCIS)	20	0 & 1000	93%	88%	n/a	n/a	1.17 x 10 <sup>-3</sup>	Guo et al. (2002)

Abbreviations: DCE = Dynamic contrast enhanced MRI, DCIS = Ductal carcinoma in-situ, NPV =

Negative predictive value, PPV =Positive predictive value, Mammo = Mammography, DWI = Diffusion weighted imaging.

In research examining the use of DWI in the differential diagnosis of breast lesions, Marini et al (2007) evaluated the sensitivity and specificity of DWI both summatively and for the specific histologically classified lesions included in their patient sample of sixty women with 81 breast lesions. They demonstrated several differences in design from those previously mentioned, acquiring the DWI sequence in the coronal plane as opposed to axially (in order to minimise phase wrapping and other artefacts), as well as excluding all lesions less than 1 cm in size on mammography and breast ultrasound.

The reason for these exclusions was not stated and introduces significant bias since MRI is able to detect lesions not identified by these two imaging sequences. These exclusions had the effect of reducing the sensitivity of both ultrasound and mammography, whilst not doing the same to MRI artificially increasing any margin of sensitivity between these two patient samples.

The pathology-specific focus of this research involved the presentation of the statistical variations between malignant tumours and, more specifically, invasive ductal vs. DCIS, which were displayed graphically. Histological grading of the malignant tumours was also summarised but unfortunately not related directly to the ADC values, which would have been beneficial in order to determine whether DWI was more sensitive to lesions with a higher nuclear grading such as intermediate and high-grade DCIS as suggested by (Kuhl, 2007).

The authors also further subcategorised their results by separating out all lesions greater than 3cm in size (sub group (A), n = 15). Of those tumours in sub group (A), differences were described between the ADC values measured at the centre and edges of the lesions. When examining the relevance of these variations with the pathology the authors noted that these reflected the pathological structure confirming necrotic-colligate centre due to the large area of the lesion and more cell packed edges.

While the sample used in this research included 81 breast lesions, it unfortunately did have a sample bias as the sample used did not include any lesions classified as being lobular carcinoma. This sample bias was disappointing as the concept of undertaking a pathology specific approach would have been especially useful in evaluating the performance of DWI if a more balanced sample had been used. The sensitivity and specificity reported in this study were 81% and 80% respectively, with the authors stating that had the thresholds been

increased a sensitivity of 100% could have been achieved with the penalty of a reduced specificity (67%). With respects to ADC distribution between lesions scatter pattern graphs were used, demonstrating significant threshold crossover and while not commented upon, showed that DWI is not useful for differentiating between either malignancies or benign lesions by histological type.

In research evaluating invasive ductal carcinoma by histological grading, Razek, Gabella, Denewer, and Nada, (2010) examined 59 female patients with confirmed invasive ductal carcinoma using b-values of 0, 200 and 400 s/mm<sup>2</sup>. The TR used for these sequences was a hefty 10000 (typically between 3000-5000) which was more than likely influenced by the number of phase encodings required to fill an imaging matrix of 512 x 256. This matrix size is significantly higher than those typically used for diffusion sequences, which from my experience are more commonly in the order of between 96 to 160 phase encodings and 156 to 256 frequency encodings in order to minimise artefacts and be clinically achievable. Artefacts involved with using this high matrix would have made the sequences more susceptible to movement artefact, chemical shift, with the long TR likely causing T2\* artefacts or shine through making ADC quantification impossible where these artefacts occurred over an area of interest. The low choice of b-values (200 & 400) was surprising with the authors describing these as a limitation of the study as lower b-values are susceptible to perfusion as well as diffusion effects. Perfusion sensitisation at lower b-values results in the sequence registering micro-circulatory movements (blood in capillaries) as well as cellular diffusion thus contaminating the signal and affecting ADC quantification.

One area of significance was the authors findings of significant ADC mean values for the different nuclear grading's of invasive ductal lesions (higher grades = lower ADC) with the authors further stating that in this work, patients with enlarged axillary lymph nodes showed significantly lower ADC values suggesting that DWI might have a role in staging patients as opposed to surgical staging. The identification of a link between ADC and pathological prognosis is interesting but probably one that needs further research, especially due to its shortcomings in sequence design.

In research examining a relatively small sample of 41 women with 65 pathologically proven lesions using DWI sequences, Park, Cha, Kang, Ihn, and Baik, (2007) detected 56 lesions using a b-value pairing of 0 and 1000 s/mm<sup>2</sup>. Of the 56 lesions detected four were classified as being benign, one of which was a cyst. Of the nine breast lesions missed, eight were

daughter nodules of a primary cancer and all were <1cm in size. It is not surprising that so many malignant lesions were missed because the DWI sequence used had a 1cm slice thickness, which is inappropriate for the intended application due to resolution and partial volume problems where lesions fall between the slices.

Park et al. (2007) did not give specificity outcomes for their research instead giving a detection rate of 86.2%, which makes it difficult to compare this with other articles. Nevertheless, while having some weaknesses, this research did provide ADC thresholds for the individual lesion types. The significance of giving ADC thresholds gives insight into whether quantitative DWI could be used to differentiate lesions by subtype. While the mean ADC thresholds given were significantly different for the different subtypes when taking into account standard deviations there remains some ADC threshold overlap between malignancies meaning that DWI is not able to correctly differentiate between them. The mean ADC thresholds given were invasive ductal carcinoma ( $0.89 \pm 0.18 \times 10^{-3} \text{ mm}^2/\text{s}$ ), DCIS ( $1.17 \pm 0.18 \times 10^{-3} \text{ mm}^2/\text{s}$ ) and benign lesions ( $1.41 \pm 0.56 \times 10^{-3} \text{ mm}^2/\text{s}$ ).

In similar research evaluating breast lesions by histological type, Woodhams, et al. (2009), examined 277 breast lesions using DWI, of which 15 were deemed to be mucinous carcinoma; 204 of the lesions were classified as being malignant while 58 were benign. Mucinous carcinoma of the breast is seen in about 1-7% of all cases of breast cancer and is characterised by having an organised and well defined cellular border surrounding large amounts of extracellular mucous (Woodhams et al, 2009). Two types of mucinous carcinoma are described: mixed and pure: pure mucinous carcinoma fits the definition already given while mixed types can contain invasive ductal components (ibid.). Previous research has established that lesions with mucinous and necrotic-colliquate centres demonstrate differing ADC values depending on where they are sampled (Marini et al., 2007).

The study by Woodhams et al. (2009) included a sample of DWI that was performed prior to DCE imaging and the sequence used had a b-value pairing of 0 and  $1500 \text{ s/mm}^2$ , which the authors justified by stating that it reduced the amount of artefact caused by what is described as T2-weighted shine through. The authors showed an inverse correlation between cellularity and ADC which demonstrates the influence of the cellular matrix on the resulting ADC values. This relationship between cellularity and ADC correlates well with the positional differences described by Marini et al. (2007).

The ADC threshold results showed that the 13 pure mucinous lesions demonstrated a mean ADC of  $1.8 \times 10^{-3} \text{ mm}^2/\text{s}$  as opposed to  $0.9 \times 10^{-3} \text{ mm}^2/\text{s}$  for malignant lesions and  $1.3 \times 10^{-3} \text{ mm}^2/\text{s}$  for benign lesions. The ADC threshold given for the two mixed mucinous lesions was  $1.2 \times 10^{-3} \text{ mm}^2/\text{s}$  demonstrating that mucinous lesions can fall outside the recognised ADC thresholds. The mean benign ADC values given by Woodhams et al. (2009) varied significantly from that given by Pereira, et al., 2009, which was  $1.14 \times 10^{-3} \text{ mm}^2/\text{s}$  for a b-value pairing of 0 and  $1000 \text{ s/mm}^2$ . One possible explanation between the differences is that Woodhams et al. excluded any lesions that were classified as being cystic in nature. This exclusion would be difficult in practice as differentiation between mixed mucinous and cystic lesions could prove problematic due to their similar ADC values.

Woodhams et al. (2009) demonstrate that mucinous breast lesions provide an exception to normal ADC thresholds. One possible area for further research could be to assess the positional ADC differences described by Marini et al. (2005) between the edge of mucinous lesions and compare these statistically with similar values at the edges of cystic lesions.

Kuroki and Nasu (2008) evaluated the inclusion of quantitative DWI to breast MRI sequences in a sample of 84 patients with both invasive and non-invasive breast cancers this time as a tool to assess response to treatment. Forty seven of the cancers sampled were lesions that had been treated with neo-adjuvant chemotherapy (NAC). Following NAC the authors stated that when evaluating DWI in conjunction with DCE sequences, specificity remained the same as DCE alone but sensitivity and negative predictive value were improved significantly (50% vs. 86%). In addition, it was demonstrated that DWI was especially useful for characterising lesions that only enhanced in later phases of DCE sequences. The importance of differentiating lesions with late contrast enhancement following NAC is that they could be either remnant cancers or fibrotic change where the NAC has been effective which exhibit similar patterns of enhancement.

Whilst DWI does allow quantitative differentiation between malignant and benign lesions of the breast it does not appear to be of any use in differentiating subtypes of lesions within those classified as being malignant. The exception to the thresholds for differentiating benign from malignant lesions is mucinous breast cancer which Woodhams et al. (2009) found had significant ADC threshold crossover. This crossover is problematic but not enough to undermine the overall performance of DWI as these lesions are relatively rare.



### *DWI as a standalone procedure without DCE imaging*

Partridge et al. (2010), examined differential diagnosis of mammographically and clinically occult breast lesions on diffusion-weighted MRI, in 91 patients with 118 breast lesions (91 benign, 12 DCIS, 15 invasive carcinoma). While using lesions detected using DCE MRI, these authors examined the use of DWI as a standalone non contrast method of detecting and characterising breast lesions. The sample of lesions in this study provides a particularly useful insight into the performance of DWI as it contained a significant proportion of patients with DCIS. As previously mentioned, Rijnsburger et al. (2010) reported this as a potential weakness of breast MRI. The sensitivity values obtained were 96% while specificity and PPV were 55% and 39% respectively. While this study intended to examine the performance of DWI as a standalone procedure, the authors themselves admitted that they were not blinded to the DCE MRI findings. They have stated that DWI shows promise for the future screening of breast lesions without contrast media, which demonstrates bias considering its specificity was only 55%, which is almost the same (56%) as quoted by Kuhl et al. (2007) for mammography alone. It appears the authors have specifically targeted lesions detected by mammography and in doing so have excluded those lesions detected by ultrasound or DCE MRI. These exclusions have increased the sensitivity of DWI in this sample as DCE MRI has significantly higher sensitivity than mammography and as such it has reduced the number of lesions identified. The effect of these exclusions has resulted in a high reported sensitivity but relatively low specificity when compared to the authors previously mentioned studies. The importance of increased sensitivity is less important than the specificity of this sample which is in the same order as mammography raising the question why bother with MRI?

In examining similar research, which also sought to detect breast cancer without DCE imaging, Suzuki, Kuroki, Nasu, Nawano, Moriyama, and Okazaki, (2007) used a combination of DWI and STIR (short tau inversion recovery) sequences. Seventy women were examined (aged 24-83) using a STIR sequence acquired axially with a DWI sequence acquired using the same spatial location at the same slice thickness (5mm). The b-value pairing used for these patients was 0 and 1000 s/mm<sup>2</sup>. These two sequences were then reported using lesion morphology and intensity only. The authors decided not to employ the use of quantitative ADC values in assessing these lesions, choosing only to examine the signal intensity associated with the DWI sequence. The findings of this research again showed a high level of sensitivity (97%) with the authors not giving specificity values. The failure of this research in giving specificity values was acknowledged by the authors and

explained as being due to the study being intended as a feasibility evaluation. While it was interesting to see a different approach used to evaluate breast DWI it is difficult to assess the credibility of this approach without the evaluation of specificity values.

The use of DWI and STIR imaging in evaluating breast cancer is not commonplace. The appeal of undertaking these sequences without conventional DCE would be to simplify breast MRI, in order to reduce the time and cost involved in order to better justify it as a screening tool. In the few studies that have been undertaken DWI and STIR sequences show mixed results with their specificity not greatly improved over that of mammography alone.

#### *b-value selection:*

Selection of b-values for DWI of the breast is a contentious issue with much of the literature differing in opinion as to the optimum b-values. In order to calculate ADC values a minimum of two b-values are required with one of the pair being or very close to 0. The use of higher b-values reduces the contribution of perfusion effects with Woodhams et al. (2011) describing how that microperfusion effects have not been observed at b-values of less than  $600 \text{ s/mm}^2$  for normal fibroglandular breast tissue, consistent with descriptions of its low vascular architecture. When determining the optimal b-values used for breast DWI several factors must be taken into account. Higher b-values show better lesion conspicuity with T2 like characteristics at the expense of SNR, while low b-values can be affected by the vascular nature of a tumour (perfusion effects) resulting in T2\* (T2 shine through) artefacts. In order to balance out these factors the approach that some researchers have undertaken is to increase the number of b-values used, taking an average of the ADC values to characterise potential lesions.

In 2009, Pereira et al. used DWI to assess breast lesions and compare the use of different b-values. This study was undertaken with the goal of determining a consensus as to how many and which b-values should be used in differentiating between benign and malignant breast lesions. The authors examined 45 patients with 52 focal breast lesions using five b-values and four separate pairings, determining sensitivity and specificity for all combinations. It was found that the statistical differences between using multiple b-values and b-value pairings did not affect the sensitivity or specificity of the examination significantly to warrant the extra time required. They did however comment that ADC values obtained using low b-values were more likely to be influenced by vascular perfusion, while high b-value combinations

resulted in less SNR (signal to noise ratio). The b-value combination of 0 and 750 s/mm<sup>2</sup> was found to be the best with a sensitivity of 92.3% and a specificity of 96.2%.

In addition, Pereira et al. examined the resolution of DWI stating that DWI can fail to categorise breast lesions because of the limited capability of recognising small lesions on the ADC maps, especially when smaller than 1cm. Exclusion criteria for this study were neo-adjuvant treatment before MRI which the authors describe as affecting the ADC values due to post treatment fibrosis. Other exclusion criteria for this study included non-mass like enhancement due to diffuse tumour spread and partial volume effect, benign cysts and lesions not visible on the DWI sequence, two of which were missed in this study (0.6 and 0.9cm).

This article helps to justify the inclusion of DWI clinically by demonstrating that it can be used in a manner that will not increase costs significantly (especially with regards to time) while greatly increasing the specificity of the examination. It also provides particularly useful recommendations as to which b-value pairing provide the best specificity and sensitivity, clarifying an area in the literature that is confusing with all authors appearing to use significantly different b-values, thus making direct comparisons of the different studies more difficult.

An earlier work published by Guo, et al. (2002) examined the sensitivity and specificity of DWI as an adjunct to DCE breast MRI. This study examined 52 patients with 55 lesions, 24 of which were benign, while the remaining 31 were classified as being malignant.

Interestingly, 10 of these patients underwent sequences of four b-values before the authors settled on 0 and 1000 s/mm<sup>2</sup>. No reason for this change in method is given, nor is there any justification given for the reasons to select 0 and 1000 s/mm<sup>2</sup>. They reported sensitivity of 93% and specificity of 88%. The authors found that they missed a 4mm DCIS lesion with this being one of the 10 patients who underwent a four b-value DWI sequence. With the benign lesions three were cysts which the authors included in calculating their ADC thresholds unlike Pereira et al. (2009) and Woodhams et al. (2009) who excluded these stating that they would artificially inflate the mean ADC values when they could be assessed without the use of MRI.

### *The effects of Gadolinium on DWI sequences:*

Gadolinium is a paramagnetic rare-earth metal with high toxicity that is used in a more chemically inert form (chelated) as an MRI contrast media to shorten the T1 relaxation time of tissues. Breast cancers are associated with angiogenesis or proliferation of blood vessels which results in both higher (three to five fold increases in parenchyma contrast uptake) and quicker contrast uptake than surrounding breast tissues. Gadolinium has a biologic half-life of between 1.5-2 hours (Carbonaro, 2011) and when undertaking DWI sequences as an adjunct to DCE imaging there is some inconsistency in the literature as to whether DWI sequences should be undertaken prior to, or following, the administration of gadolinium and what relevance this may have on ADC thresholds.

The effects of gadolinium upon DWI sequences were investigated by Rubesova et al. (2006). This study included a patient cohort of 77 patients with 110 breast lesions with DWI sequences being undertaken in all but five cases prior to contrast administration. The authors wanted to assess the effect gadolinium might have on ADC thresholds, which they identified as being controversial due to presumed T2\* (gradient echo) changes being responsible for increases in mean ADC thresholds over and above those used without gadolinium. This theory of T2\* effects increasing the ADC values was tested by evaluating seven patients both pre and post-contrast (gadolinium) stating that no such increases were seen. But this needs to be considered cautiously since there were only seven lesions included. Overall, they found specificity of 86%, but as in some other studies, they excluded cysts from the benign lesions. These authors raise the issue of resolution of DWI and use 0.7cm as the value below which they suggest DWI. This potential unreliability is due to the resolution problems associated with its low imaging matrix (less pixel numbers) and partial volume effects by using slice thicknesses over 4mm. The image matrix selected for use on breast DWI sequences is balanced by the need to image quickly in order to prevent motion artefacts, while the rationale for a slice thickness of around 4mm is related to the need for imaging a large area (whole breast) again within time constraints. Another subject addressed by these same authors in this research was the classification of fibroadenomas with DWI, which in this sample were all correctly classified as being benign, further stating that since fibroadenomas often mimic malignant tumours due to their contrast enhancement, that DWI would be a good tool to supplement DCE imaging in assessing these lesions. Two lesions that were misclassified in this sample were one lobular invasive carcinoma and one lobular carcinoma in-situ. The lobular carcinoma in-situ was described as having a bland histological structure

similar to that of benign structures, while no reasons were given for missing the invasive carcinoma. Information regarding that particular lesion's size would have been useful in determining whether its absence was due to resolution related volume effects as discussed by (Pereira, et al. 2009).

Further research investigating the administration of gadolinium before undertaking DWI sequences of the breast was undertaken by Ogura, A., Hayakawa, K., Miyati, T. And Maeda, F. (2008) in a non-clinical environment using phantoms. They found that gadolinium had no effect on the magnetic susceptibility of DWI sequences, disagreeing with the theory expressed by Rubesova et al. (2006) regarding presumed T2\* (gradient echo) changes. The downside of this research is that because it was undertaken in a purely scientific environment it does not completely disprove the influence of physiological causes of gadolinium influencing ADC values. Ogura et al. (2008) discuss one such physiological theory of how gadolinium could affect ADC by causing blood vessel shrinkage and changes to blood viscosity.

In a small clinical study involving 19 women, Yuen et al. (2009) directly investigated the physiological effects of gadolinium on ADC values in imaging breast carcinomas, confirming the previous work about the vascular effects on ADC values discussed by Ogura et al (2008). They found that when using a b-value pairing of 0 and 1000 s/mm<sup>2</sup> the administration of gadolinium actually reduced the mean ADC readings by 23%, the importance of which the authors suggested was due to the negation of perfusion effects on the quantification of breast lesions. These effects were deemed to be beneficial as they emphasised the 'pure-cellularity and malignant potential of breast carcinoma.

In evaluating the research on imaging breast carcinoma with DWI it appears evident that gadolinium does not influence the magnetic susceptibility and is potentially useful in overcoming those problems with perfusion described by Peters et al. (2010).

When examining the effects of gadolinium in reducing perfusion effects it also gives validity that b-value pairings less than 1000 s/mm<sup>2</sup> would be more appropriate post contrast. The research by Yuen et al. (2009) highlights the importance of using relevant ADC thresholds derived from research involving gadolinium.

Overall the literature shows overall the administration of contrast media does in fact affect ADC thresholds. These effects on DWI appear to increase mean ADC values by up to 23%

(Yuen et al. 2009). The significance of these affects upon ADC thresholds seems to be a positive one as it negates perfusion effects, potentially allowing the use of lower b-values, which allows better signal to noise ratios (SNR). The most important theme emerging from the literature is that any ADC thresholds used need to be consistent with their clinical use being either pre or post-contrast administration.

### *High field strength breast DWI*

The next step for breast DWI appears to be using higher field strengths such as 3T (Tesla) which is available for clinical use in small numbers in New Zealand. Whether higher field strengths improve or degrade these sequences clinically is what I am seeking to clarify in evaluating this literature.

An article by Peters et al. (2010) discusses b-value selection for DWI sequences of the breast using a 3 Tesla magnet. In this article the authors examined 73 patients with 90 lesions using a combination of five b-values (0, 150, 499 and 1500 s/mm<sup>2</sup>).

Perfusion effects associated with using low b-values were described as improving with 3T scanners due to larger signal to noise ratios allowing the use of higher b-values clinically without affecting the quality of the scans. Interestingly, despite this discussion upon b-values the authors stated that while better SNR was evident in using 3 Tesla when compared to 1.5 Tesla, that because of the problems associated with fat saturation and shimming the larger magnetic field inhomogeneities of 3 Tesla scanners, that in reality DWI obtained at higher field strengths as in this study demonstrated results no better than those previously published using 1.5 Tesla.

In research also examining the use of 3T and DWI in evaluating breast lesions, Lo et al. (2009) evaluated a relatively small cohort of 31 female patients with suspected breast lesions on mammography and ultrasound. All of these lesions had tissue samples taken with the histology results used as gold standard. These authors give sensitivity and specificity values of 91 and 90% respectively, which is relatively high compared to the lower values (63%) quoted by Yabuuchi, et al. (2008). This research does not however mention any of the problems with fat saturation and shimming that Peters et al. (2010) discuss in their article. Lo et al. (2009) undertook this research using a b-value pairing of 0 and 1000 s/mm<sup>2</sup> with the authors justifying their choice by suggesting that by using a relatively high b-value that the sequence is less susceptible to the motion of capillary blood vessels. The use of a 3T MRI

scanner in this case appears to have allowed these authors the luxury of a significantly superior SNR overcoming the detrimental effects that Pereira et al. (2009) described as being a shortcoming for this b-value pairing.

These two articles have significantly different findings with one stating that the higher SNR's achieved with 3T scanners is negated by overcoming the associated artefacts with the other giving the opposite opinion, quoting high sensitivity and specificity values and failing to mention any of the shortcomings expressed by the other. This is obviously an area which needs further research.

#### Spectroscopy:

The themes identified in the literature regarding the clinical use of breast spectroscopy are: sensitivity and specificity, the presence of and composition of identifying a choline peak; the use of high field scanners; and the use of spectroscopy in determining the effectiveness of neo-adjunct chemotherapy.

The principles or origins of MRI spectroscopy come from a technique widely used in the scientific analysis of chemicals known as nuclear magnetic resonance (NMR) spectroscopy. Breast spectroscopy is a newer concept in MRI and is primarily used to differentiate between benign and malignant lesions.

Breast spectroscopy is normally undertaken in the clinical setting (1.5 Tesla) using single voxel techniques. These single voxel spectroscopy techniques ( $^1\text{H}$ ) can differ technically using either the STEAM (stimulated echo acquisition mode) or PRESS (point resolved spectroscopy sequence) methods of spatial registration and voxel interpolation (Wang, 2008). When using single-voxel techniques it is the presence and position of metabolites in the spectra that can be quantified (ppm) and not the amount or concentration. When seeking to quantify metabolites, the literature suggests that these require additional spectroscopic imaging, which is more accurate when utilising larger field strength MRI scanners. These techniques typically involve multiple-voxel (Chemical shift imaging (CSI)) sequences such as those used in neurological spectroscopy.

### *Sensitivity and specificity*

As the human breast consists of mostly adipose tissue this poses problems for breast spectroscopy. Fat saturation techniques are already quite robust for breast MRI but what does pose a problem is fat or adipose tissue within the area of interest (within the voxel). Fat contamination of the MRI spectra is a tricky, problem as the voxel is a cube, while lesions are typically asymmetrical. Signal from fat selected within a voxel can be nullified by the use of saturation bands, where this fails it can result in reduced metabolite amplitude. Bartella et al (2007) discuss this effect describing it as being lipid side banding. Lipid side banding is defined as being caused by mobile lipids within the selected voxel that result in spurious echoes, which interfere with visualisation of the total choline (tChol) peak. Voxel size can be changed in order to fit within a selected lesion for mass type lesions, with the penalty of doing this being an exponential increase in scanning time increasing the probability of patient movement artefact.

In research that examined single proton spectroscopy alongside DCE and DWI sequences, Tozaki and Fukama (2009) evaluated 165 patients with 171 highly suspicious (BI-RADS 4 and 5) lesions. These lesions were further classified as being 124 mass and 44 non-mass lesions. Of these lesions, pathology confirmed malignancy in 91 cases and benign results in 80 cases. The overall sensitivity and specificity of the breast spectroscopy sample was 44% and 85% respectively compared to 97% sensitivity and 67% specificity for the same sample using DWI alone. These authors used a  $15\text{mm}^3$  voxel for their spectroscopy stating that smaller lesions were poorly defined (less than 15mm), while DWI imaging was useful for characterising lesions of any size. Breast spectroscopy in this study performed poorly (44%) with the authors putting this down to the large number of non-mass lesions (predominantly DCIS 25/44), which when excluded (lesions less than  $15\text{mm}^3$ ) increased the sensitivity to 82% while the specificity decreased to 69%.



**Table 2: Comparative sensitivity and specificity breast imaging table (Mammography vs. Breast MRI with and without DWI and Spectroscopy sequences).**

	No. Of Malignant lesions	No. Of Benign lesions	b-Value mm <sup>2</sup> /s	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Thresholds mm <sup>2</sup> /s	Authors
<b>Mammo</b>	167 DCIS  89 high grade DCIS	n/a	n/a	56%  52%	n/a	55%	n/a	n/a	Kuhl (2007)
<b>MRI DCE</b>	167 DCIS  89 high grade DCIS	n/a	n/a	92%  98%	n/a	59%	n/a	n/a	Kuhl (2007)
	30	20	n/a	100%	62.5%	n/a	n/a	n/a	Huang et al. (2004)
<b>DWI</b>	29	87	0 & 600	96%	56%	39%	98%	1.6 x 10 <sup>-3</sup>	Partridge et al. (2010)
<b>DWI&amp;DCE</b>	61	14	0, 500 & 1000	92%	86%	97%	71%	1.1 x 10 <sup>-3</sup>	Yabuuchi et al. (2008)
	26 (2 DCIS)	26	0 & 750	92.3%	96.2%	n/a	n/a	1.24 x 10 <sup>-3</sup>	Pereira et al. (2009)
	31 (1 DCIS)	20	0 & 1000	93%	88%	n/a	n/a	1.17 x 10 <sup>-3</sup>	Guo et al. (2002)
<b>Spectro &amp; perfusion</b>	18	12	n/a	100%	87%	82%	n/a	n/a	Huang et al. (2004)
<b>Spectro</b>	31	26	n/a	100%	88%	91%	n/a	n/a	Bartella et al. (2004)

Abbreviations: DCE = Dynamic contrast enhanced MRI, DCIS = Ductal carcinoma in-situ, NPV = Negative predictive value, PPV =Positive predictive value, Mammo = Mammography, DWI = Diffusion weighted imaging, Spectro = MRI breast spectroscopy.

In further research investigating enhancing non-mass lesions in the breast with spectroscopy, Bartella et al. (2007) quoted 100% sensitivity and 85% specificity in assessing these lesions using a relatively modest sample of 32 patients. The most encouraging findings of this study were that the authors stated that positive predictive value was increased from 20% to 63%; the relevance being that biopsy could have been avoided for 17 of 25 cancers where those lesions displayed a choline peak on spectroscopy.

Tozaki and Fukama (2009) contradict the results achieved by Bartella et al. (2007) finding that spectroscopy is not reliable for assessing non-mass like lesions or more specifically DCIS. The authors state that in their study false negative results were shown in 68% of DCIS cases.

The main differences between these studies were the number of patients with DCIS (2 vs. 25 respectively) and the spectroscopy voxel size, which Bartella et al. changed to suit each patient depending upon lesion size ( $9 - 25\text{mm}^3$ ) while Tozaki and Fukama used a  $15\text{mm}^3$  voxel for all patients sampled (165 lesions, BI-RADS 4 and 5). The significance of undertaking spectroscopy with smaller voxels means that lesions are more likely to be a better fit within those dimensions reducing lipid contamination and flattening out of any choline peaks.

In additional research evaluating the performance of spectroscopy for 46 lesions (19 malignant, 21 benign), Tse et al. (2003) demonstrated 89% sensitivity and 100% specificity when assessing breast lesions greater than 1.5cm in size. The reported specificity for this research was 100% with no false positive results, while the sensitivity was 89% with two lesions displaying false negative results. For these two false negative results the authors reported that both displayed absent HER2 results. This relationship was further explored in research by the same principle authors Tse, Yeung, King, Cheung and Yang (2007) who explain this mechanism by describing how in carcinogenesis, HER2 encodes for EFGR (epidermal growth factor receptor) which results in abnormal persistent activation of tyrosine kinase activity that triggers mitoses even in the absence of EFGR. Breast cancer lesions are typically tested for HER2 status by immunohistochemistry as cancers with positive results are likely to be more aggressive. The reason that HER2 receptor status is tested in breast cancer is because patients with positive results are candidates for treatment with therapeutic antibodies such as trastuzumab (Herceptin) Tozaki and Hoshi (2010).

Research that investigated the use of pharmacokinetic or contrast kinetics in correctly diagnosing false positive spectroscopy results was that by Geraghty et al. (2008). In this study these authors evaluated a relatively small sample of 16 pathology confirmed invasive ductal carcinoma. The results of this study found that 14/16 of these lesions were correctly diagnosed as being malignant and because no benign lesions were included in this study the authors gave a positive predictive value of 88% as opposed to sensitivity and specificity. Geraghty et al. (2008) stated that because choline levels may not be as elevated in some breast cancers as they are in others, they are sometimes not detected at all resulting in false-negative diagnoses. The hypothesis in this research was that the two false negatives in this sample could be positively diagnosed by conventional DCE contrast kinetics and a pharmacokinetic ratio of intracellular and extracellular exchange rates labelled  $k_{21}$ .

Breast lesions with DCIS can present in two forms: pure, in which the lesion contains only encapsulated disease, or mixed with other components such as invasive disease. Pure DCIS can be described as a non-mass like lesion as it is contained the milk ducts. Tse, Yeung, King, Cheung, Yang (2007) reported that all of the six pure DCIS lesions evaluated in their study demonstrated negative spectroscopy, speculating that this was in relation to the less aggressive behaviour of the disease. The false negative findings of Tse et al. (2007) are consistent with findings of Tozaki and Fukama (2009) who also found problems differentiating between DCIS and more specifically non-mass lesions. This problem with differentiating non-mass lesions exists primarily due to the fact that these lesions have length (contained within the ducts) rather than volume and in most cases prove a poor fit for a spectroscopy voxel.

### *Choline*

With breast spectroscopy the presence of a choline peak is used to determine the pathological nature of the sample. Current clinical spectroscopy sequences at 1.5T (single voxel) typically involve visual identification (qualitative assessment) of a choline peak or threshold SNR levels (quantitative) in order to differentiate areas of malignancy from areas that may otherwise contain benign breast tissue. When assessing breast tissue it is increased levels in the total amount of the metabolite choline (tCho) that may be a reflection of increased membrane turnover by rapidly replicating cells in active breast cancers (Bolan, Nelson, Yee & Garwood, 2005).

In the earliest investigation of breast spectroscopy Roebuck et al. (1998) identified the potential of identifying tCho as a biomarker of malignancy of breast cancers. The alterations in the metabolite concentrations of breast cancers at 1.5 Tesla are discussed in more detail by Katz-Brull, Lavin & Lenkinski (2002) who describe increased content of phosphomonoesters, phosphodiesteres as well as water soluble choline metabolites (phosphocholine and glycerophosphocholine), which these authors describe as contributing to composite choline levels. Benign breast tissues are described as having lower levels of phosphomonoesters and phosphodiesteres as well as nondetectable levels of the composite choline signal. These authors then examined five previously published studies involving detection of composite choline levels, achieving a pooled analysis of up to 92% sensitivity and specificity.

While tChol is identified in malignant breast cancers some researchers such as Stanwell et al. (2005) have identified choline as being present in some healthy volunteers and lactating women, which they use as reasoning to further interrogate the choline spectrum in order to prevent false positive results that would effect specificity. Bartella et al. (2004) identify the metabolite phosphocholine as being the largest component contributing to the Choline peak from cancerous breast tissue. Total choline (tChol) peaks that indicate malignancy occur between 3.14 and 3.34ppm (Sardenelli, 2009), while the specific phosphocholine peak occurs between 3.22-3.23 ppm Tozaki & Fukuma, 2009).

Tozaki and Fukuma (2009) discuss a secondary choline peak at 3.27-3.28 ppm which they say is assigned to glycerophosphocholine, taurine, and myo-inositol and can be defined as being benign. Because of the identification of this extra benign peak Tozaki and Fukuma appear to have broken the metabolite spectrum into a further subcategories, whereas Roebuck et al. (1998) and Katz-Brull et al. (2002) discussed a single marker for malignancy or tChol at 3.2ppm. The consequences of incorporating these secondary and potentially benign metabolites into the tChol could then be perceived as having the potential to reduce the overall specificity of breast spectroscopy because if malignancy is absent and a peak is observed around these frequencies this could result in false positive results. Stanwell and Mountford (2007) also discuss this resonance at 3.28 ppm which they describe as being present in healthy volunteers further stating that the term tChol is a misnomer when the resonance is centred at 3.28ppm as this could be a benign result and not one associated with malignancy.

In a review of the literature by Haddadin et al. (2009) there is some clarification of the issues previously identified in determining which metabolites are contained within the tChol peak at 3.2ppm. Haddadin et al. (2009) describe how the tChol peak at 3.2ppm is used as the cancer biomarker for in-vivo studies, these metabolites can currently only be separated and quantified during ex-vivo and in vitro studies. The relevance of quantifying the tChol peak was given because higher levels of phosphocholine concentrations have been shown to be present with higher grade tumours and because of the need for quantitative measurements when accurately comparing research.

Bolan et al. (2003) discuss how further analysis of the composite choline spectrum can be undertaken using spectral fitting software. This research was however undertaken at 4 Tesla using a purely research based MRI scanner.

#### *Effect of contrast on spectroscopy*

Breast spectroscopy is more commonly undertaken post-contrast administration. This occurs due to the need to identify lesions effectively and as such the effect if any that Gadolinium has on the choline spectrum needs to be assessed.

In a review of the six most commonly available gadolinium based contrast media, Lenkinski, Elian, Wang & Goldberg (2009) evaluated their effect on both phantoms and rats implanted with breast adenocarcinoma. This research was undertaken using a 3 Tesla MRI scanner with mixed results. The six contrast agents were subcategorised by their ionic charge (either neutrally or negatively). The negatively charged agents Magnevist, Multihance and Dotarem had the effect of both broadening the line width and reducing the height (up to 40%) of the choline spectra, while the neutrally charged contrast media (Omniscan, Optimark and ProHance) had little or no effect on the spectra. These authors then concluded by recommending that neutral gadolinium chelates be used in assessing breast cancers with spectroscopy.

In similar research that evaluates spectroscopy at 1.5 Tesla Tozaki & Maruyama (2010) undertook research on 30 patients with two different gadolinium based contrast media. These authors referred directly to the work by Lenkinski et al. using one negatively charged gadolinium chelate (Magnevist), and one neutral chelate (Omniscan) on all of the 30 patients comparing spectra obtained on separate days. The findings of this research found no statistically significant changes in either line width or spectra height at 1.5 Tesla.

### *High-field breast spectroscopy and future developments*

The themes emerging in the literature assessing spectroscopy at higher field strengths such as 3T are increased spectral resolution, generation of larger SNR, metabolite quantification and the use of multi-voxel techniques in assessing breast lesions.

Sensitivity and specificity of breast spectroscopy is increased at higher field strengths with Haddadin et al. (2009) describing that tChol could be more consistently detected in smaller lesions at 3T than at 1.5T. These increases in specificity are possible because of the larger SNR generated by these scanners allowing the use of smaller voxels that are a better fit for evaluating smaller tumours, as well as better shimming which allows a reduction in the lipid contamination within the voxel. Higher field spectroscopy also allows better identification of the individual metabolites by increased spectral sampling resolution (Stanwell & Mountford, 2007). These metabolites exist at the same location (ppm) on the spectrum and this literature review has shown that increased field strengths appear to further break down the atomic coupling of these compounds allowing better peak differentiation, which at 3T makes it possible to identify and quantify the different choline peaks at 3.2 and 3.28ppm.

Multiple voxel spectroscopy techniques such as chemical shift imaging (CSI) and spectroscopic imaging (SI) can be undertaken using 1.5T MRI scanners in either 2 or 3 dimensions allowing capture of quantitative metabolite information. Metabolite quantification itself can be undertaken in two ways: SNR thresholds, or by absolute quantification giving values to each peak. Quantitative single voxel breast spectroscopy at 1.5T is questionable as choline observed at this field strength is known to contain both benign and malignant signals. Multiple voxel techniques at higher field strengths such as 3T allows choline differentiation due to better spatial resolution and has the added advantage of being able to acquire larger areas of breast tissue without incurring time penalties, allowing comparisons to be made between both normal breast tissue and potentially malignant lesions. Multiple voxel spectroscopy can also allow the inclusion of multiple lesions should these fall within the area being sampled.

An example of CSI MR Spectroscopy research at 1.5T was undertaken by Su (2008) who examined 36 malignant and 9 benign lesions using SNR thresholds to quantify choline. The sensitivity given for CSI spectroscopy at 1.5T was 81% while the specificity was 78% using a

sequence that took 15 minutes. While sensitivity and specificity of this sequence appeared to be high, Su (2008) suggested that while regional choline levels were observable that diagnosis upon thresholds was in fact difficult due to lipid contamination. Very similar research undertaken on a 3T scanner using a 3D spectroscopic sequence to evaluate 32 malignant and 12 benign lesions gave both higher sensitivity (97%) and improved specificity of 84% using a CSI sequence that took approximately 11 minutes (Gruber et al. 2011). The authors of this work stated that 3T MR spectroscopic imaging showed significant differences in SNR between low-grade and high-grade tumours, which they declare has not previously been demonstrated. The authors of this study concluded by stating that 3D spectroscopic sequences yield both high diagnostic sensitivity and specificity allowing the study of heterogeneous and multi-centric breast tumours. These increases can help to simplify acquisition planning as spectra can be generated post acquisition amongst any of the voxels within the volume.

An example of single voxel quantitative spectroscopy at high-field strength is that by Bolan et al. (2003) who undertook research using a 4T MR scanner. These voxels were then quantitatively analysed again by using SNR thresholds in order to determine tChol levels. This research was one of the earlier spectroscopy studies at higher field strengths and involved 105 women, 86 of whom were presenting with suspicious mammography diagnosed lesions with the remainder being followed up for response to chemotherapy. Bolan et al. (2003) found that for the lesions sampled they only achieved 46% sensitivity to tChol levels. This sensitivity was poor and the authors did not give specificity values perhaps in part due to the patient sample to which they did not give a breakdown of lesion types (mass or non-mass) only stating that half were pathology confirmed. What Bolan et al. did state was that tChol levels did not appear to be related to different histological types of cancer which is in direct contradiction to Gruber et al. (2011) who undertook a similar study with what was a multi-voxel sequence.

In research that involved the use of a 7T MRI scanner, Klomp et al. (2010) examined four patients; three of these patients had pathology confirmed breast cancer with the fourth being a healthy volunteer. This study involved both single and multiple voxel spectroscopy with the authors using resolution which they state was comparable to positron emission tomography (PET) scanning, which involves ionising radiation.

Two of the four patients examined by Klomp et al. (2010) went on to undergo chemotherapy which was then able to be followed up with quantitative data for comparison. The spectrums achieved by this scanner allowed assessment of both phospholipid metabolite levels, energy metabolite levels (inorganic phosphates, adenosine and glycerol diesters) and pH levels which, are calculated by using the chemical shift effect.

The relevance of detecting energy metabolites means that MR spectroscopy can provide functional data such as that provided by PET scanners. While not readily available for clinical use and with this research only involving four patients, high field MR scanners such as this 7T appear to show great promise for the future of MR spectroscopy of the breast, especially in the follow up of a patient's response to chemotherapy.

#### Summary:

In evaluating the literature assessing the performance of DWI sequences in breast MRI the most common themes that emerge are increases in PPV and specificity that are the more important factors and not sensitivity, which appears to be in the ninetieth percentile when DCE is evaluated as a standalone procedure.

There remain variations in sensitivity and more particularly specificity in the literature with some like Yabuuchi et al. (2008) giving a range of values. Individual studies sensitivity and specificity values appear to be heavily dependent upon method and lesion type, while most researchers attempt to control for external validity there are so many factors involved (research design, equipment, lesion types and reviewer experience) that no two articles appear to give the same results and direct comparison is difficult.

While there have been studies evaluating DWI without DCE imaging it is evident that this is at the sacrifice of sensitivity to lesions smaller than 1cm.

There appears to be much debate as to the selection of b-values for DWI sequences. Pereira et al. (2009) recommended the use of 0 and 750 s/mm<sup>2</sup>, using similar scanning parameters and equipment to those which have been used in the sample that I am evaluating. In evaluating the differences in specificity and PPV by study type, b-value choice and ADC threshold selection there appears to be no standardisation between any of the studies and again because of similarities of design the same ADC thresholds given by Pereira et al. (2009) will be used in this research.



When evaluating breast DWI undertaken using higher field strengths (1.5 vs. 3 Tesla) it appears that this is better due to the problems of any higher signal to noise ratios being negated due to the fat saturation difficulties associated with shimming what are larger field in-homogeneities (Peters et al. 2010).

One theme emerging from the literature is the description of lobular carcinoma in-situ (LCIS) being a bland histological structure with lower ADC thresholds than other malignancies. This and the potential pitfalls in differentiating mucinous carcinoma from benign lesions is a theory that could be further evaluated in this research by sub analysis of the results by lesion type and ADC value.

Following evaluation of the literature involving DWI it is apparent that it has a role to play as an adjunct to DCE imaging rather than as a standalone procedure. When used as an adjunct to DCE imaging DWI sequences positively influence both the specificity and PPV of the examination.

When evaluating spectroscopy techniques current lesion specificity does appear promising for large mass type lesions (>15mm), while lesion specificity of non-mass type lesions such as pure DCIS appears to need further evaluation due to lower levels of sampling in the literature.

When assessing the sensitivity and specificity of spectroscopy the relationship of HER2 status and false negative spectroscopy results as reported by Tse, Yeung, King, Cheung, Yang (2007) is an interesting one, which is also worthy of further investigation.

Current clinical practice when using MR spectroscopy is by use of  $^1\text{H}$  or single proton spectroscopy, which allows visualisation of choline metabolites when a sampled tissue is malignant. When reading the literature surrounding the choline peak it becomes clear that this peak is made up of several compounds all of which result in what is termed the tChol or total choline peak. Several researchers such as Tozaki and Fukama (2009) discuss a secondary choline peak at 3.28ppm which could be present in benign tissues. This secondary peak is significantly smaller than that demonstrated by malignant lesions with this metabolite present between 3.2 and 3.26ppm. At present with 1.5T MRI scanners confirmation of results is by either (qualitative) visualisation of the choline peak between 3.2 and 3.26ppm or (quantitative) SNR thresholds. As larger field MRI scanners become available, the future of breast spectroscopy appears to be using multiple voxel spectroscopy sequences with these

techniques allowing quantification of metabolites, as well as comparisons between normal breast tissue and those containing lesions. High field breast MR spectroscopy also shows promise in evaluating a patient's response to chemotherapy both by metabolite quantification and qualitative visual assessment of any choline peaks.

The effects of contrast upon spectroscopy at 1.5 Tesla appear to be negligible, at where higher field strengths are used current research suggests that neutrally bound gadolinium chelates have less effect on choline spectra.

## Study methods

### *Setting and Patients*

This study involved a retrospective audit of the case records for women who had undergone breast MRI at one unit (Radiology Department) over a two year period between 01/12/2009 and 30/11/2011. This research was undertaken ethically with departmental approval complying with the Unitec research and ethics regulations, following the approval of a research proposal.

To be included, patients needed to have undergone dynamic contrast enhanced (DCE) MRI along with diffusion-weighted sequences (DWI) and/or spectroscopy. Patients fulfilling these criteria were then required to have matching pathology results. The pathology results were considered to be the gold standard to which the MRI results were compared.

Of all of the patients fulfilling these criteria, DCE and DWI imaging was performed on 52 female patients aged between 23 and 78 years old, with a combination referral diagnosis of suspicious lesions (17 patients), pre-op query multifocal disease (23 patients), and difficult diagnosis by other imaging modalities (12 patients). In total this sample included 68 pathology confirmed breast lesions; DCIS (n = 11) IDC (n = 30), Lobular (n = 15), Mucinous or with necrosis (n = 4) and 8 lesions considered to be benign.

The spectroscopy sample used in this research included 21 lesions from 21 separate patients, 18 of these lesions were confirmed by pathology as being malignant: IDC (n = 11), Lobular (n = 5), Mucinous (n = 2). The remaining three lesions in this sample were confirmed by pathology as being benign.

All lesions identified on DCE breast MRI sequences were cross referenced by this researcher for location and lesion type with clinically available pathology results. Lesions were first classified as being benign (0) or malignant (1) before being further classified by pathological subtype such as ductal or lobular carcinoma in-situ, invasive ductal and mucinous or necrotic type lesions. Pathology results considered relevant were those from biopsies either fine needle or radiology guided as well as surgical wide local excisions, lumpectomies and mastectomies where no chemo or radiotherapy had occurred pre-procedure (see exclusion criteria).

### *Exclusion criteria*

From the DWI sample patients that have undergone radiotherapy prior to MRI were excluded, along with those not undergoing routine protocols. Radiotherapy can cause fibrosis and scarring at a cellular level altering the DWI and spectroscopy results which is beyond the scope of this study. Patients who had moved during their examination were excluded due to issues with diagnostic accuracy. Movement during DWI scans is not always noticed upon completion of the scan and is better appreciated as artefact on the ADC map, causing these to be immeasurable.

The exclusion criteria for spectroscopy differed from those used with DWI. While examinations that demonstrate patient movement will also be excluded since this movement changes the location of the voxel used, the main exclusion criteria for spectroscopy was that all lesions being sampled were greater than  $1\text{cm}^3$ . Voxel size is a limitation of the MRI scanner and the spectroscopy sequence being used.

### *Imaging Parameters*

All imaging was undertaken using a General Electric (GE) 1.5 Tesla Signa HDx MRI scanner using a dedicated 8ch breast coil, with images analysed on a commercially available CADstream™ (version 4.1) workstation. The protocols used for breast MRI included Axial STIR (Short Tau Inversion Recovery) of both breasts, VIBRANT™ (Volume Imaging for BREast AssessmeNT) imaging of both breasts acquired in the axial plane, incorporating a mask phase, before the use of 10mls of MultiHance® (Gadobenate Dimeglumine) gadolinium based contrast media administered intravenously by an automated power injector and following a short delay of 40 seconds five further VIBRANT™ phases acquired before subtraction is undertaken.

DWI sequences were acquired using a 36cm field of view (FOV),  $96 \times 160$  matrix sampled with 16 signal averages (NEX) using a dual spin echo EPI sequence with a parallel imaging factor of 2 (ASSET®). The TR used was 2800ms with TE optimised ~ 80 ms. A b-value pairing of 0 and  $750 \text{ s/mm}^2$  was used with gradients in the superior and inferior directions. Slice thickness was 5mm with a 1mm gap taking 3 minutes and 58 seconds.

Parallel imaging techniques such as General Electric's ASSET® work by using spatial encoding from multiple RF detector coils to reduce the number of phase-encoding steps required to fill k-space. This reduction in phase-encoding steps primarily speeds up

acquisition times and is discussed by Woodhams et al (2011) who described how ASSET® when used in conjunction with DWI EPI sequences allowed reductions in image distortion, susceptibility artefacts.

VIBRANT™ sequences were acquired using a 320 x 320 matrix, optimised for enhancement/temporal resolution (less than 90 seconds) acquiring a pre-contrast mask followed by five phases post-contrast administration. Scanning breast tissue in less than ninety seconds is the key to breast MRI as after approximately two minutes normal breast tissue begins to enhance following contrast administration. FOV for the VIBRANT™ sequences was dependent upon breast size and undertaken using either 32 or 35cm.

Spectroscopy was undertaken using a commercially available single voxel sequence (BREASE™) with a TR of 2000ms and TE of 155ms, a line width of less than 18 was required for all sequences. No less than eight saturation bands were placed for each lesion with these encroaching into the voxel where intra-voxel fat was observed. Spectroscopy sequences were optimised to fit the lesions with 20mm<sup>3</sup> (32 NEX), taking 4 minutes 48 seconds; 18mm<sup>3</sup> (56 NEX), 8 minutes; 15mm<sup>3</sup> (83 NEX), 11 minutes 20 seconds.

The remaining post contrast sequences include high resolution T2-weighted sagittal imaging of both breasts individually. DWI sequences were acquired post-gadolinium in the axial plane using the same imaging locations as the STIR sequence.

Spectroscopy sequences were performed if lesions are greater than 1 cm in circumference and time allowed.

All sequences with the exception of the axial STIR and spectroscopy protocols were imaged using a manually selected fat saturation technique that requires the MRT to confirm or optimise the position of the fat peak on the NMR spectrum.

ADC maps were generated retrospectively by post processing the DWI sequences using commercially available software (FunctTool™, GE healthcare, Buckinghamshire, United Kingdom). The audit involved 64 pathology confirmed lesions from 43 women. All lesions were visualised using CADstream™ software, with correlation of STIR, DWI and ADC maps before manual placement of a region of interest (ROI) to generate quantitative ADC data.

All imaging was stored digitally as per hospital policy in digital format using a picture archiving and communications system (PACS) Sectra medical systems, Linköping, Sweden.

All data was also stored on optical disc as per departmental protocols which, allowed easier identification and retrieval for retrospective analysis.

### *Data analysis*

Following evaluation of the data by way of a retrospective audit all data was separated into subgroups depending upon the pathology results:

1. Malignant e.g. (IDC, Lobular, Mucinous, DCIS).
2. Benign.

This data was tabled with outcomes being positive (1) or negative values (0) in subgroups containing the separate tests i.e. Pathology, DWI and Spectroscopy. These tables were then evaluated statistically using freely available software ([www.vassarstats.net](http://www.vassarstats.net)) allowing sensitivity and specificity values to be generated.

Sensitivity and specificity are key concepts in understanding a health tests performance to both positive and negative results. Positive predictive value (PPV) and negative predictive value (NPV) are also ways by which procedures are evaluated.

Sensitivity was used in this research to demonstrate MRI's (DCE, DWI and spectroscopy's ability to correctly identify those patients breast disease.

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

(Lalkehen and McCluskey, 2008)

Specificity was used in this research alongside sensitivity in order to evaluate the ability of these same methods to correctly identify those patients without disease.

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}}$$

(Lalkehen and McCluskey, 2008)

Like sensitivity and specificity, PPV and NPV are often used to evaluate clinical tests. PPV is often used in clinical research in substitution of specificity in order to determine an

examinations likelihood of predicting the presence of disease where the examination gives a positive result.

$$\text{Positive predictive value} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

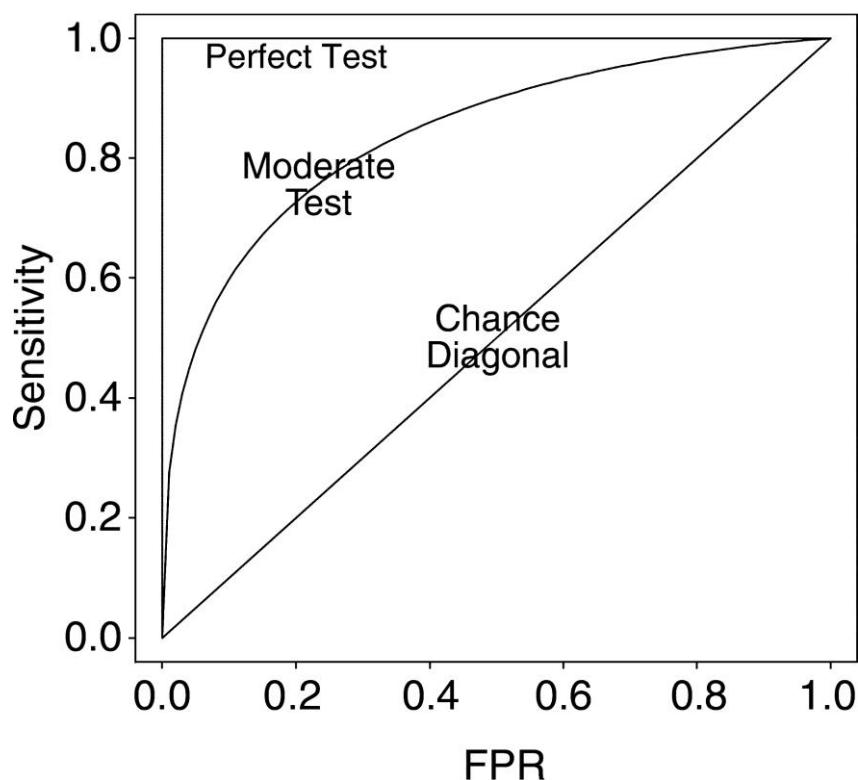
$$\text{Negative predictive value} = \frac{\text{True negatives}}{\text{True negatives} + \text{False negatives}}$$

(Lalkehen and McCluskey, 2008)

This data was also evaluated using further medical statistical software ([www.medcalc.org](http://www.medcalc.org)).

This program allowed generation of receiver operator curves (ROC) and a scatter pattern graph demonstrating b-value thresholds.

**Figure 5: ROC curve example**



ROC curve demonstrating the relationship of sensitivity vs. FPR in a graphic form.

(Obuchowski, 2003)

While sensitivity and specificity values are to be given, receiver operating characteristic (ROC) curves will allow visual demonstration of these variables taking into account both true

and false positive results. ROC curves are a measure of true positive rates versus false positive rates or sensitivity versus 1 – specificity (Obuchowski, 2003).

## Results

### Dynamic Contrast Enhanced (DCE):

Of the 68 lesions, DCE imaging alone demonstrated a sensitivity and specificity of 90% and 62.5% respectively (Vasser Stats, accessed 11/05/2012).

For this data set sensitivity gave 95% confidence interval (95%CI) values of between 78% and 95%, while 95%CI for the specificity calculations was between 25% and 89%. Positive predictive value (PPV) was 83% (95%CI: 72-91%).

The DCE sample included 3 false positives which included lesions with chronic inflammatory change, duct hyperplasia and fibrotic changes. The false negatives (6) in this sample were evenly divided between 3 lesions containing DCIS (intermediate grade) and 3 lesions of mixed IDC and DCIS pathology.

### DWI

Of the 68 lesions, DWI imaging alone demonstrated a sensitivity and specificity of 98% (95%CI: 90-100%) and 63% (95%CI: 26-89%) respectively. Positive predictive value (PPV) was 95% (95%CI: 81-96%), (Vasser Stats, accessed 14/06/2013).

The DWI sample included one false negatives, which involved a lobular lesion. This lesion had an ADC value just outside the thresholds used for classifying malignancy. The false positives (n=3) in this sample included lesions with ductal hyperplasia, chronic inflammatory and fibrotic changes.

### DCE and DWI:

In combining the results between these two sequences there first needs to be a rule for when the two methods are in disagreement.

When examining the coded results a value of 0 is given for any result that is negative in relation to the pathology results, while a value of 1 is given to those results that are positive.



When both DCE and DWI are in agreement  $1 + 1 = 1$  and  $0 + 0 = 0$

When there is disagreement there are two possible outcomes

a)  $1 + 0 = 0 + 1 = 1$

b)  $1 + 0 = 0 + 1 = 0$

When applying rule a) Sensitivity achieved is 100% (95% CI: 92-100%) while specificity is 38% (95% CI: 10-74%), (Vasser Stats, accessed 14/06/2013).

Using this rule there are no false negatives and 5 false positives.

When applying rule b) Sensitivity achieved is 90% (95% CI: 79-96%) while specificity is also 88% (95% CI: 47-99%), (Vasser Stats, accessed 14/06/2013).

Using rule b) there are 6 false negatives and 1 false positive.

In the sample of DCE and DWI imaging combined there was one lesion that tested false positive on both sequences with this being classified as being chronic inflammatory change.

**Table 4: Audit data (Pathology, DCE & DWI)**

HISTOLOGY	ADC (10 <sup>-3</sup> mm <sup>2</sup> /s)	DCE	DWI	HISTOLOGY	ADC (10 <sup>-3</sup> mm <sup>2</sup> /s)	DCE	DWI
DCIS	1.13	0	1	IDC	1	1	1
DCIS	0.98	1	1	LOBULAR	0.93	1	1
DCIS	1.11	1	1	LOBULAR	0.76	1	1
DCIS	1.15	1	1	LOBULAR	0.74	1	1
DCIS	1.13	0	1	LOBULAR	1.15	1	1
DCIS	0.79	0	1	LOBULAR	0.87	1	1
DCIS	0.98	1	1	LOBULAR	0.82	1	1
DCIS	1.1	1	1	LOBULAR	0.99	1	1
DCIS	0.99	1	1	LOBULAR	0.98	1	1
DCIS	0.99	1	1	LOBULAR	1.3	1	0
DCIS	1.16	1	1	LOBULAR	1.06	1	1
IDC	0.79	1	1	LOBULAR	1.24	1	1
IDC	0.73	1	1	LOBULAR	0.645	1	1
IDC	1	1	1	LOBULAR	1.16	1	1
IDC	1	1	1	LOBULAR	0.9	1	1
IDC	1.06	1	1	LOBULAR	0.9	1	1
IDC	0.67	1	1	MUCINOUS	0.98	1	1
IDC	0.82	1	1	MUCINOUS	1.12	1	1
IDC	0.93	1	1	MUCINOUS	1	1	1
IDC	1.13	1	1	MUCINOUS	1	1	1
IDC	0.79	1	1	NEG PATH	1.5	1	0
IDC	0.95	1	1	NEG PATH	1.11	0	1
IDC	0.9	1	1	NEG PATH	0.93	0	1
IDC	1.16	1	1	NEG PATH	1.03	1	1
IDC	1.22	1	1	NEG PATH	1.31	1	0
IDC	1.1	1	1	BENIGN	1.54	0	0
IDC	0.99	0	1	BENIGN	1.6	0	0
IDC	1.16	0	1	BENIGN	1.45	0	0
IDC	1.11	0	1	n = 68			
IDC	1.08	1	1	Mean = 1.01985			
IDC	1.06	1	1	n = 57			
IDC	0.83	1	1	n = 62			
IDC	0.86	1	1	Key			
IDC	0.76	1	1	1 = Positive			
IDC	1	1	1	0 = negative			
IDC	0.99	1	1				
IDC	0.99	1	1				
IDC	0.98	1	1				
IDC	0.86	1	1				
IDC	0.89	1	1				

## Spectroscopy:

Spectroscopy ( $^1\text{H}$ ) was performed on 21 patients all of which were included in the sample undergoing both DCE and DWI imaging. The spectroscopy sample included 19 lesions which were confirmed by pathology as being malignant (IDC (n = 11), lobular (n = 5), mucinous/colloid (n = 2). This sample also included three lesions confirmed by pathology as being benign, one of which giving a false positive spectroscopy result.

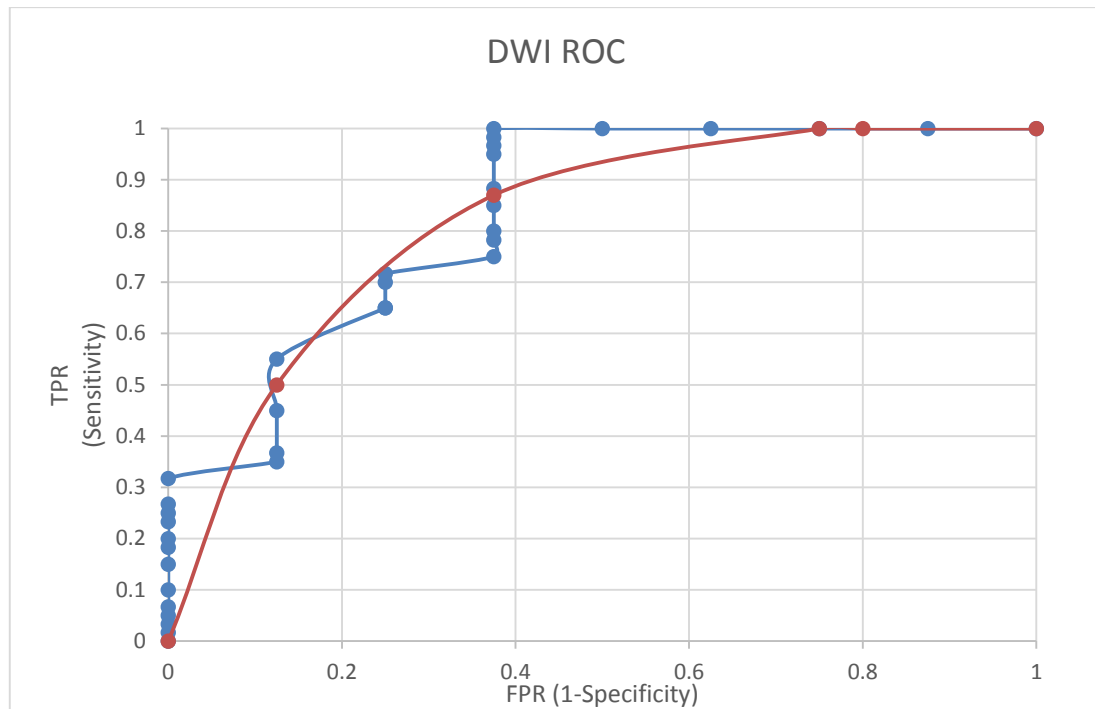
The Sensitivity for the spectroscopy sample was 100% (95%CI: 78-100%). Unfortunately due to insufficient numbers of benign lesions any specificity calculations would prove unreliable. Positive predictive value (PPV) was 90% (95%CI: 68-98%), (Vasser Stats, accessed 24/05/2012).

**Table 5: Audit data (Pathology, spectroscopy)**

Histology	Spectroscopy
IDC +DCIS	1
IDC	1
IDC+DCIS	1
IDC+DCIS	1
IDC	1
IDC+DCIS	1
Mucinous (COLLOID)+DCIS	1
IDC+DCIS	1
LOB	1
LOB	1
IDC+DCIS	1
IDC+DCIS	1
LOB	1
LOB	1
IDC+DCIS	1
IDC(NECROSIS)+DCIS	1
IDC	1
LOB	1
Benign	1
Benign	0
Benign	0
n = 21	n =19

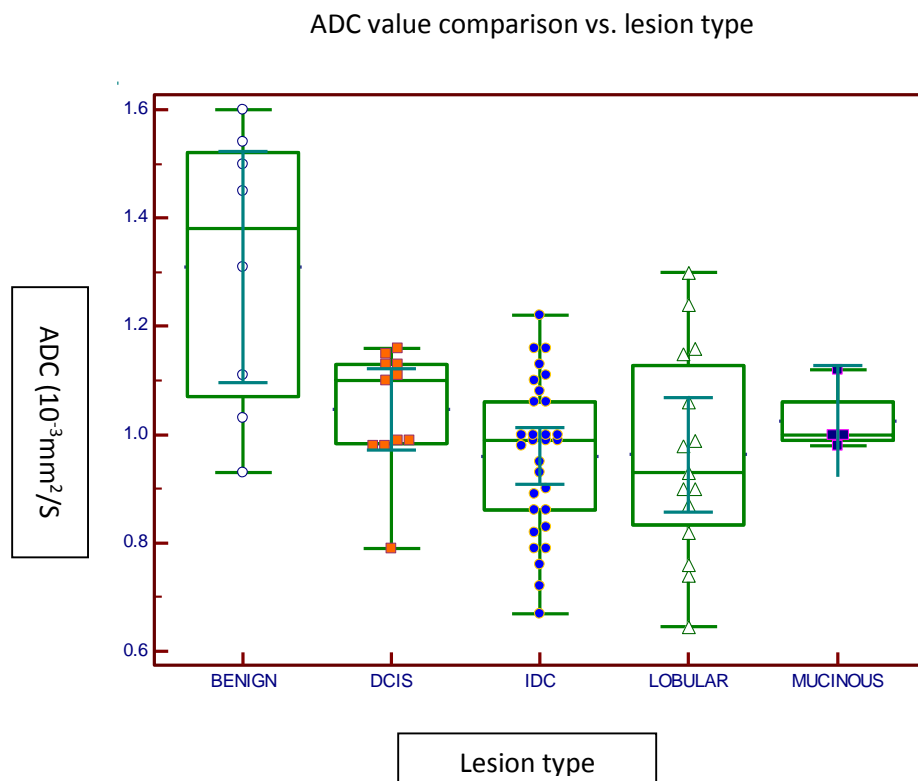
Key 1 = condition present, 0 = condition absent

## DWI ROC



The DWI ROC curve gives an area under the blue empirically plotted curve of 0.844 (95% CI: 73-92). The blue ROC curve is empirically fitted demonstrating the relationship between ADC thresholds and the associated pathology. The red line is fitted demonstrating a smoother more classically looking ROC curve. Evaluation of the red line also shows how higher sensitivities are possible for DWI at the expense of specificity.

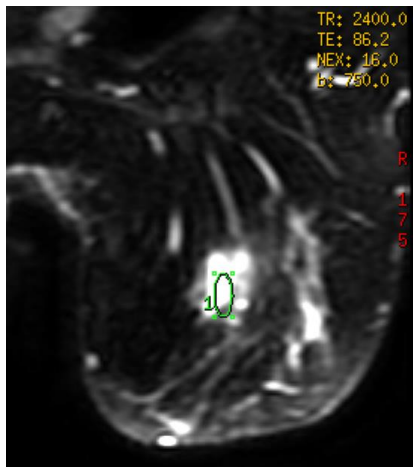
## ADC Comparison



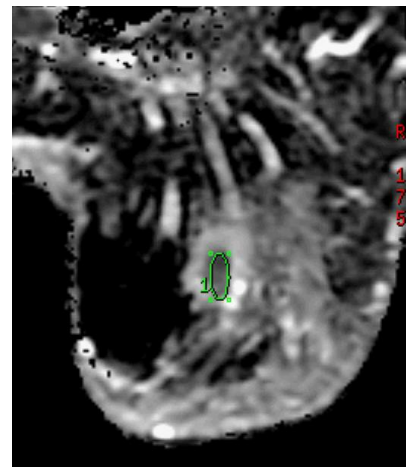
A comparison graph demonstrates the distribution of ADC values between the different tissue confirmed pathologies. The ADC thresholds used in this study were values equal or lower than  $1.24 \times 10^{-3}\text{mm}^2/\text{s}$  for malignancy. Benign classification was given to all lesions with values greater than  $1.25 \times 10^{-3}$ , which were the same thresholds as used by Pereira et al. 2009 when using this b-value pairing. This graph demonstrates that the malignant lesions are randomly grouped within the malignant thresholds with no specific pattern of distribution evident that would help to differentiate between the different lesion types. When examining mucinous lesions the mean ADC of  $1 \times 10^{-3}\text{mm}^2/\text{s}$  demonstrated by this study was outside those given by Woodhams et al. (2009) of  $1.81 \times 10^{-3}\text{mm}^2/\text{s}$  for pure mucinous lesions and  $1.21 \times 10^{-3}\text{mm}^2/\text{s}$  for mixed mucinous lesions.

### Example 1: Chronic inflammatory change. (False positive)

A 46 year old female with pathology confirmed left breast cancer that underwent breast MRI with the clinical question being query multifocal/bilateral disease. MRI demonstrated a suspicious right breast lesion with positive DCE and DWI results. Following a bilateral mastectomy this right sided lesion was found to be benign demonstrating a pathology result that was deemed to be chronic inflammatory change.



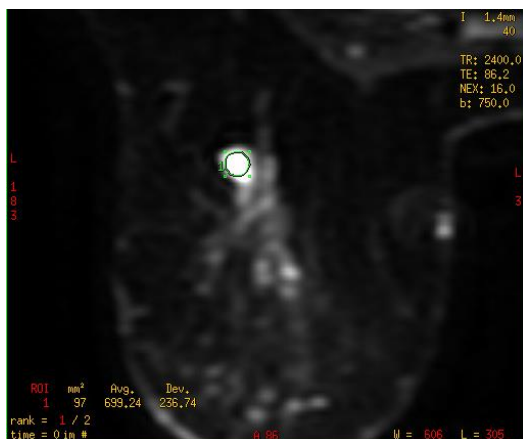
A) DWI



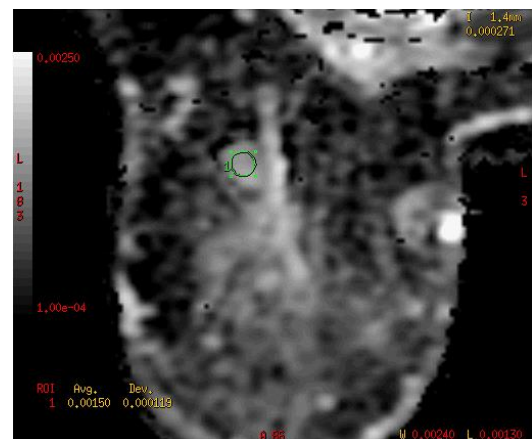
B)  $ADC = 1.03 \times 10^{-3} \text{mm}^2/\text{s}$ .

### Example 2: Benign fibrotic change (False positive).

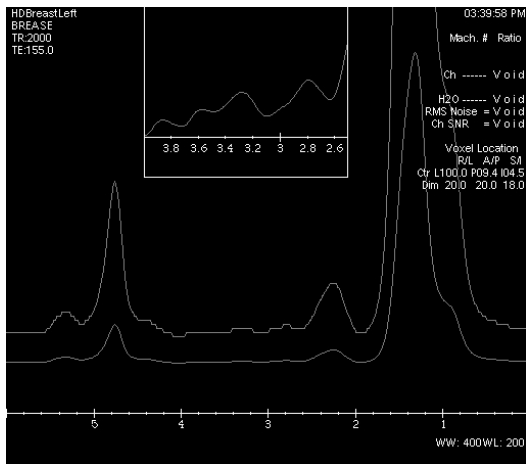
A 55 year old female with a false positive DCE and Spectroscopy result. This lesion was confirmed by pathology to be benign fibrotic change.



A) DWI



B)  $ADC = 1.5 \times 10^{-3} \text{mm}^2/\text{s}$ .

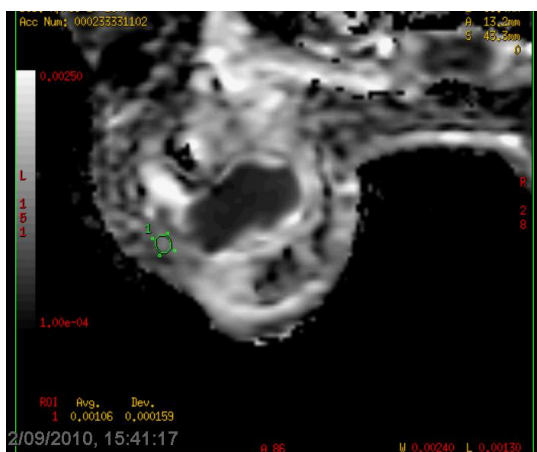


C) Choline peak visible at approximately 3.26ppm.

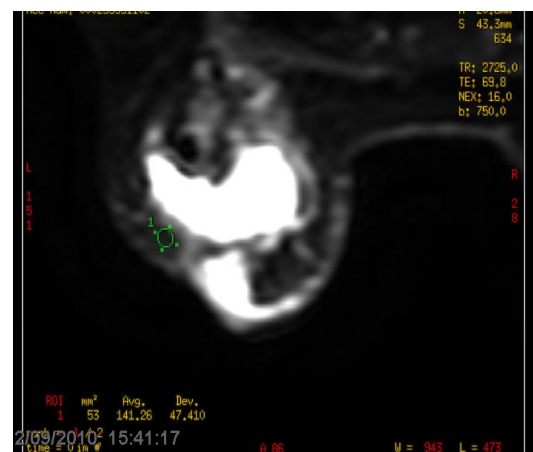
Inset = magnified portion of spectrum between 2.6-3.8ppm

### Example 3: Ductal Carcinoma In-Situ (DCIS) (False negative DCE).

A 67 year old female patient who underwent a breast MRI following a wide local excision of a pathology confirmed left breast lesion. This surgery resulted in a pathology sample with positive margins for DCIS and an MRI was requested prior to further surgery and before the patient underwent chemotherapy. DCE imaging of this breast provided a negative result, while DWI gave a malignant ADC of  $1.06 \times 10^{-3} \text{mm}^2/\text{s}$ . Pathology of the area confirmed residual intermediate grade DCIS.



A)  $\text{ADC} = 1.06 \times 10^{-3} \text{mm}^2/\text{s}$ .



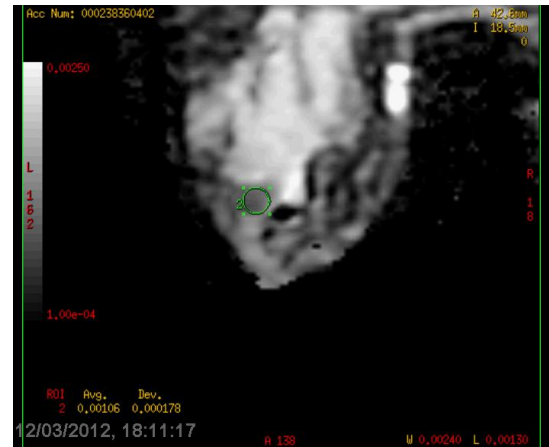
B) DWI.

#### Example 4: Lobular carcinoma.

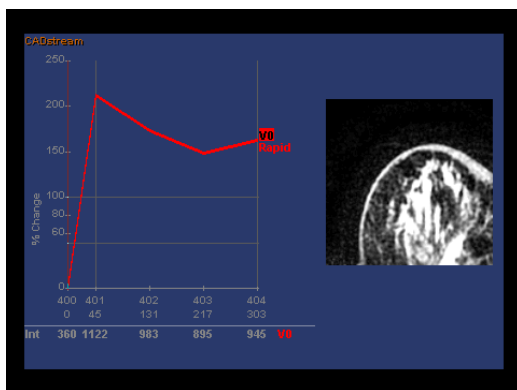
A 61 year old female with 2 MRI confirmed lesions (DCE and DWI). Pathology confirmed the MRI findings with both lesions shown to be lobular in nature.



A) DWI



B)  $ADC = 1.06 \times 10^{-3} \text{mm}^2/\text{s}$ .

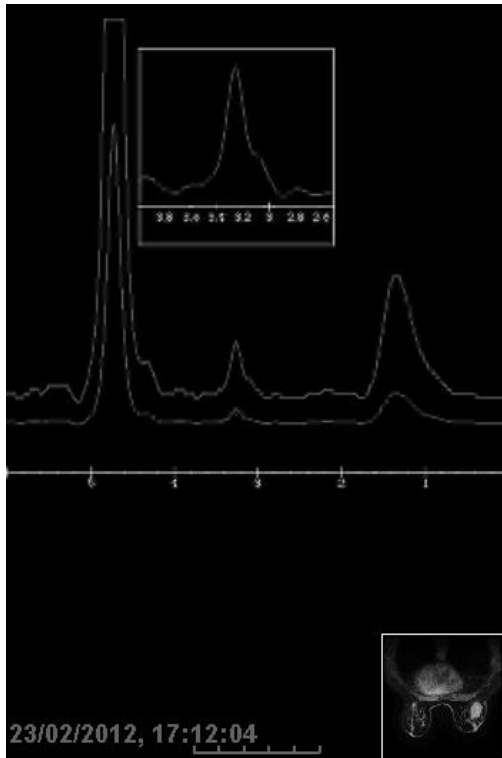


C) DCE contrast enhancement map (Malignant pattern)



### Example 5: Intra Ductal Carcinoma (IDC).

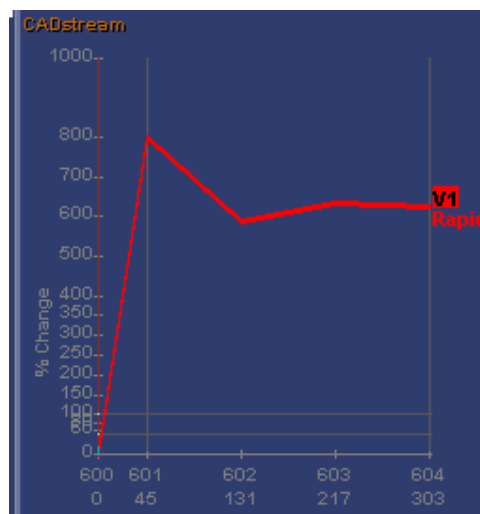
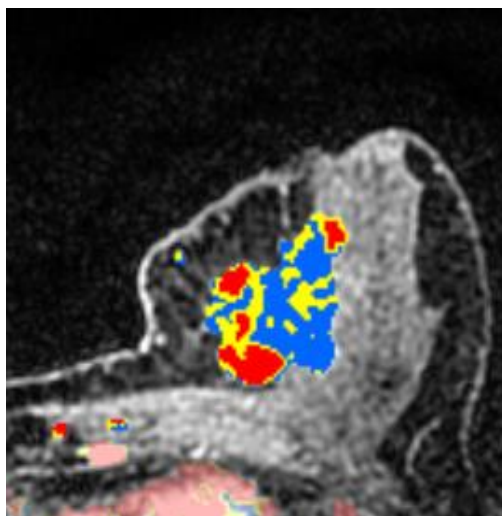
A 64 year old female with pathology confirmed IDC, which was positively confirmed as being malignant on DCE, DWI and spectroscopy.



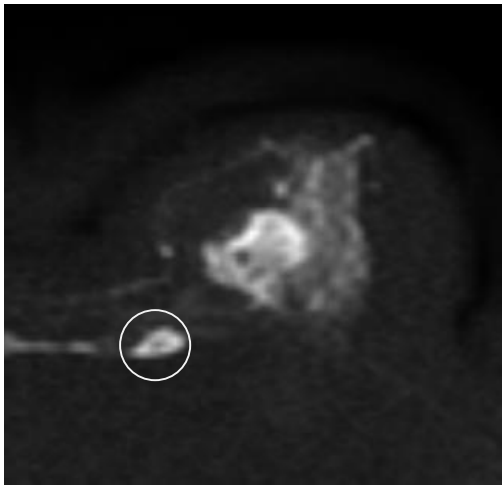
A) Spectroscopy with choline peak at 3.26ppm    B)  $ADC = 0.73 \times 10^{-3} \text{mm}^2/\text{s}$ .

### Example 6: IDC

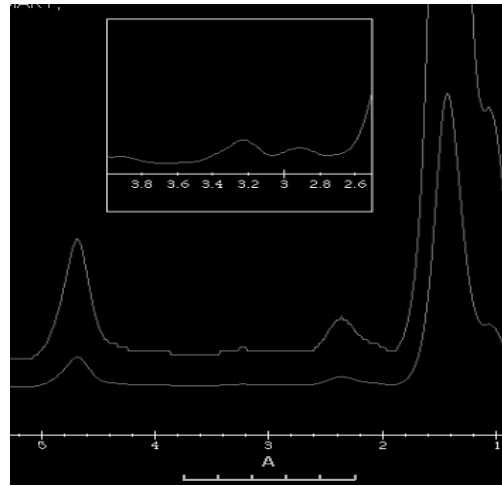
A 61 year old female diagnosed with biopsy and following the MRI pathology proven invasive ductal carcinoma.



A) CADstream™ DCE (colour overlay)    B) Contrast kinetic graph



C) DWI



D) Spectroscopy with choline peak @ 3.23ppm

This example demonstrates a clearly defined invasive ductal carcinoma which displayed a malignant contrast enhancement pattern and choline peak on spectroscopy. The ADC for this lesion was also deemed to be malignant ( $1.1 \times 10^{-3} \text{mm}^2/\text{s}$ ).

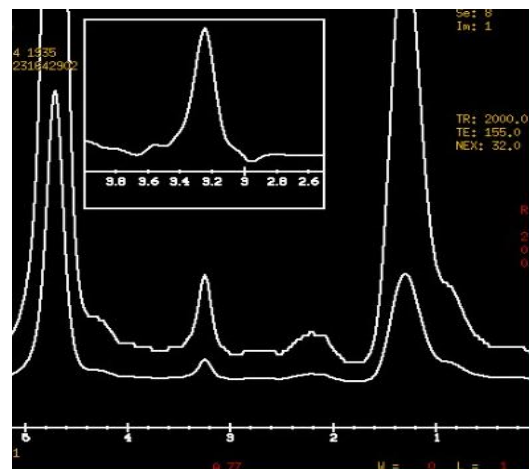
Demonstrated on the DWI (within highlighted circle) was an enhancing internal mammary lymph node not previously observed on the DCE imaging alone. This lymph node was removed and examined with the subsequent pathology confirming lymph node invasion.

### Example 7: IDC

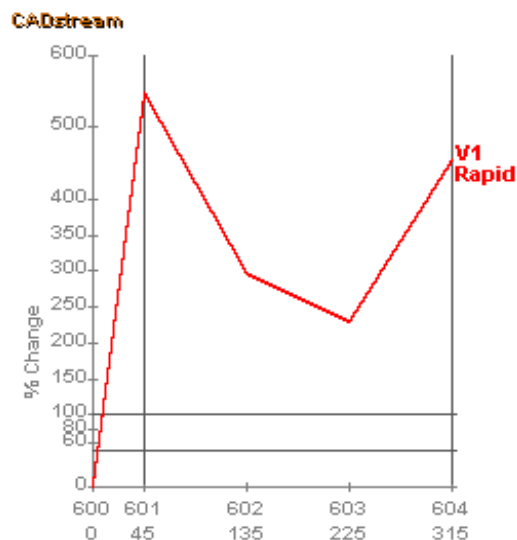
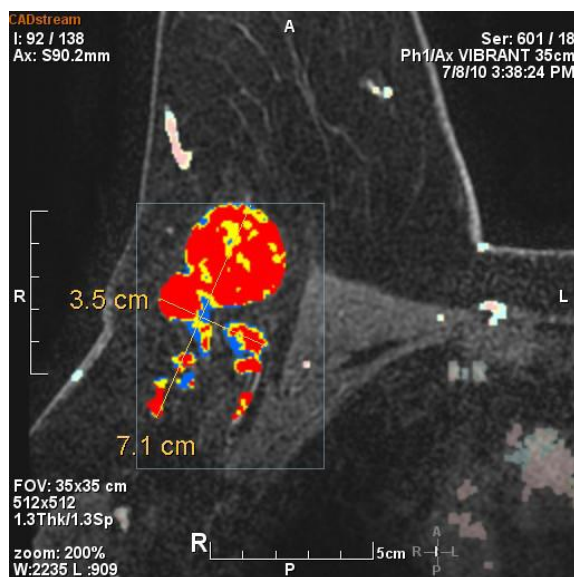
A 60 year old female thought to have a metastatic axillary lymph node following a positive biopsy. The clinical question for this patients MRI was query location of a primary breast cancer.



A) Voxel placement ( $20\text{mm}^3$ )

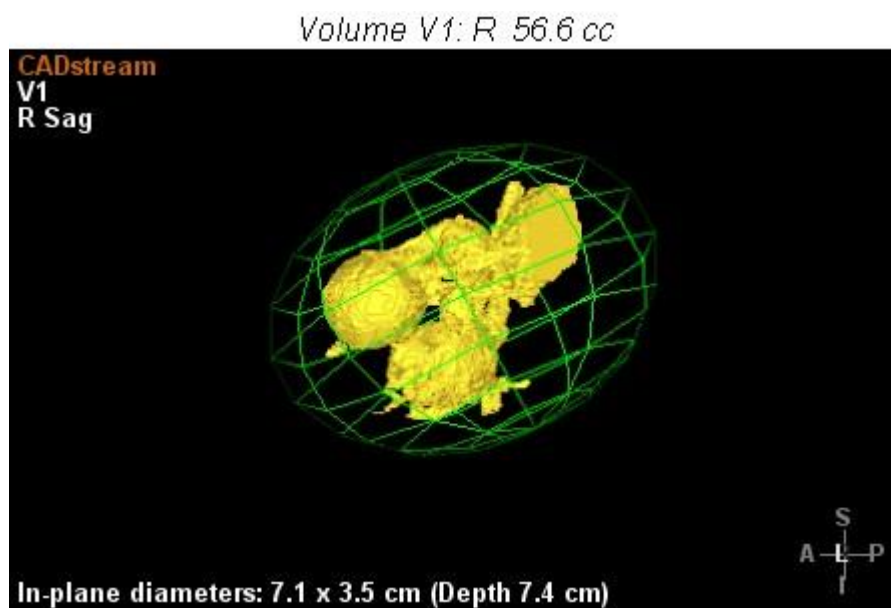


B) Spectroscopy with choline peak @ 3.26ppm



C) DCE (CADstream with colour overlay) D) Contrast enhancement graph

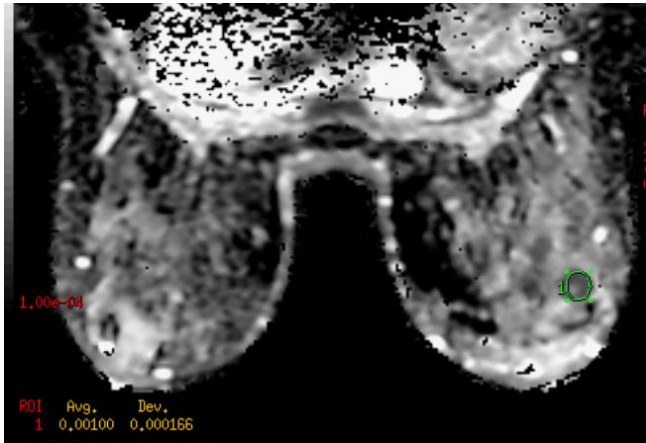
CADstream™ DCE reformat showing colour overlay and volume measurements (C) with matching contrast kinetic graph (D) showing a malignant pattern of rapid contrast uptake (over 500%) and incomplete washout. The result of the MRI was that the structure thought to be a metastatic lymph node was in fact the primary cancer demonstrated to be a complex tumour much larger than represented by both mammography and ultrasound (See 3D volume E).



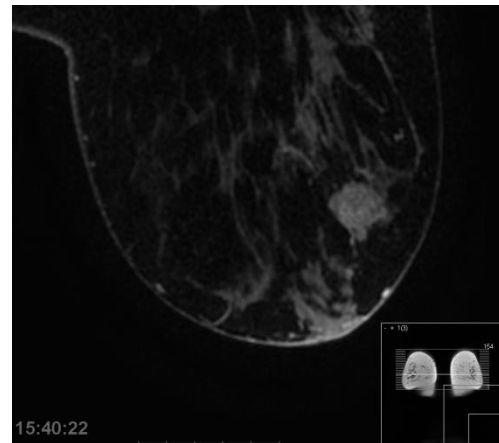
E) CADstream™ 3D volume reformat

### Example 8: IDC

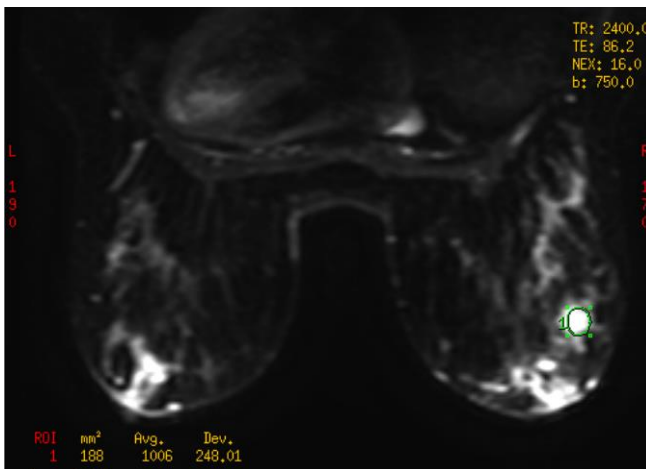
A 42 year old female patient presenting to MRI with biopsy proven IDC. The clinical question for this MRI was query extent of disease with DCE, DWI, and spectroscopy confirming the presence of this cancer, whilst excluding multifocal and bilateral disease giving this patient the choice of breast conserving surgery.



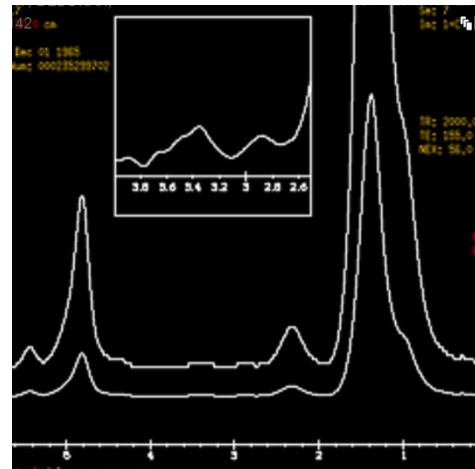
A)  $ADC=1.0 \times 10^{-3} \text{mm}^2/\text{s}$



B) DCE



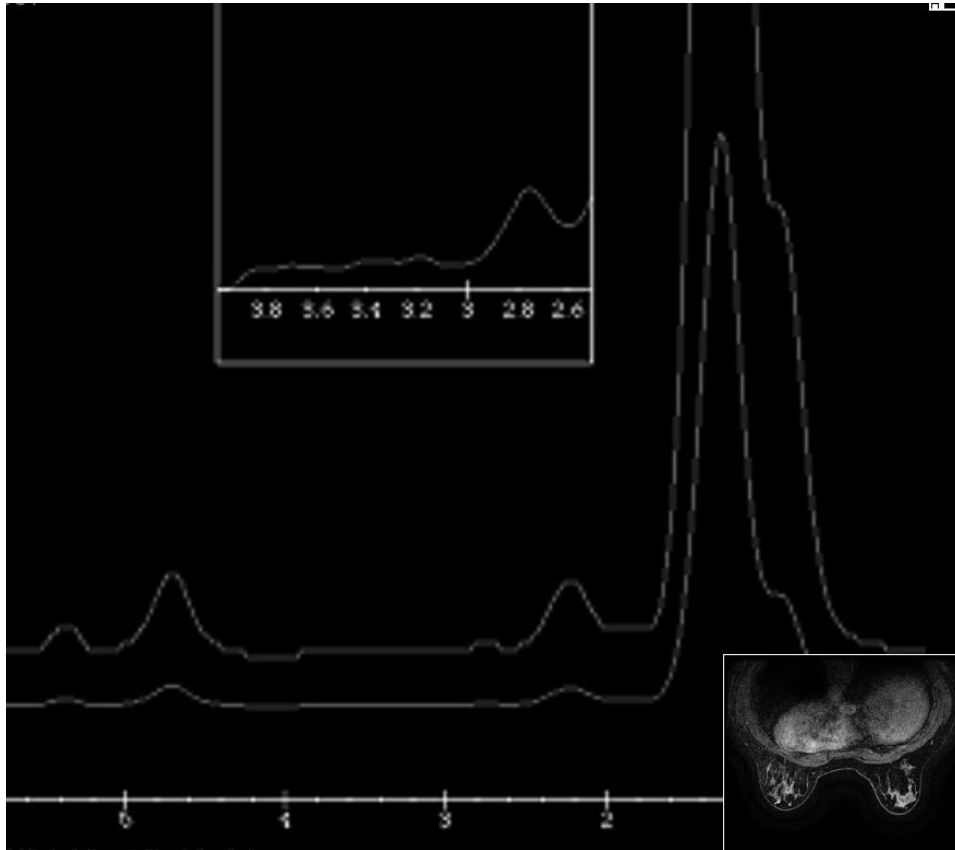
C) DWI



D) Spectroscopy with choline peak at approx 3.33ppm

### Example 9: Papilloma

A 34 year old female with a history of juvenile papillomatosis requiring yearly MRI scans. Biopsy of suspected papilloma resulted in pathology that stated no malignancy was seen and sample contained an area with a benign foreign body type reaction.



A) Spectroscopy with no visible choline peak.

## Discussion

### *Current practice*

MRI of the breast is a relatively new and expensive technique available to clinicians in order to improve the detection of breast cancer. Traditionally breast screening programs have used techniques such as clinical examination, mammography and ultrasound to identify and interrogate breast lesions. Because of its high sensitivity to breast cancers MRI has emerged as a worthy adjunct to these methods giving surgeons further options in evaluating potential breast cancers that may otherwise be missed (Yoo, 2010).

Breast MRI is more commonly used in New Zealand for evaluating women at high risk for breast cancer, such as those women with BRCA 1 and BRCA 2 gene mutations. This screening of high risk women is accepted practice and was recommended in a report published by the New Zealand health advisory (NZHTA) unit on behalf of the New Zealand Ministry of Health (Davidson, 2007). This report involved a meta-analysis of a large portion of the then current research into screening these patients.

Breast screening guidelines for the management of early breast cancer in New Zealand were published by the New Zealand guidelines group of behalf of the Ministry of Health (2009) (NZGG, 2009). These guidelines include imaging, surgical and treatment recommendations and are endorsed by the Royal Australasian College of Surgeons (RANZCS), the Royal Australian and New Zealand College of Radiologists (RANZCR) as well as the New Zealand Cancer Society. This guideline includes assessment of the relevant literature by members of both colleges making recommendations as to best practice. Both of these papers are guidelines and not legislation and, where accessible, breast MRI is typically used at the service's discretion. As both guidelines were at least three years old at the time of reading a decision was made to evaluate the current literature in order to get an insight into what other centres or countries consider normal practice and whether new technologies or applications of breast MRI exist.

### *Clinical indications for the use of breast MRI*

The New Zealand guidelines for the use of breast MRI are almost identical to those of the American College of Radiologists (ACR) whose recommendations include: evaluation of contra-lateral breast, extent of disease (multi-focality, multi-centricity), scar tissue vs. recurrence, identifying primary lesions where the axilla is involved, lesion response to

neoadjuvant therapy, screening patients with BRCA-1 and -2 Gene mutations, non-calcified disease DCIS, and imaging young patients with dense breast tissues and invasive ductal carcinoma (Raza, 2010). The ACR breast MRI guidelines can be used in three ways:

1. Preoperatively
2. Screening of women with an increased risk of breast cancer by recurrence and genetic mutation.
3. Surveillance of women with an increased risk of breast cancer by recurrence and genetic mutation.

The use of breast MRI for preoperative screening has been examined in the recently published Comparative Effectiveness of MR Imaging in Breast Cancer (COMICE) trial (Campbell, 2010). The authors of this study sought to examine the arguments for and against pre-operative breast MRI by the way of an audit of approximately 1600 patients. The COMICE trial examined the relationship between those women undergoing pre-operative breast MRI with those requiring additional surgery and concluded that there was no advantage of undertaking routine pre-operative breast MRI.

However, Dr Elizabeth Morris, a professor of radiology specialising in breast MRI, argued against the findings of the COMICE trial stating that while the sample contained two groups (MRI and non-MRI) they were selected randomly and not matched by staging and/or types of cancer (Morris, 2010). Also, the MRI sample used was also multi-centric and did not take into account the equipment (MRI scanner) or radiologist experience. The remaining and perhaps strongest criticism by Professor Morris was that this trial did not involve any patient recurrence data (*ibid.*).

The COMICE study has received a lot of publicity in New Zealand, giving ammunition to those doubting the usefulness of breast MRI for routine preoperative planning. Associate Professor Ian Campbell (2010) discusses the findings of the COMICE trial from the perspective of a New Zealand breast surgeon; he agrees in part with the study describing that while routine preoperative screening may not be useful as shown by the COMICE trial, MRI still has a role to play in assessing women with invasive lobular cancers: suspected multi-focal disease; very dense breasts; very strong family history of breast cancer; or some discordance in findings between triple assessment modalities (clinical exam/mammography and ultrasound).

MRI of the breast is an expensive examination in comparison to the previously mentioned modalities and the increasing workloads pressures on clinical scanners often poses problems with clinical availability. At current capacity it would be very difficult to accommodate all patients requiring routine preoperative breast MRI if this were a requirement in clinical guidelines. Both the investment in imaging equipment and the staffing costs are higher for MRI compared to mammography; breast MRI currently costs about 6 times more than mammography in private practice and is likely to be similarly different in the public hospital system.

Breast MRI has been available for at least seven years at my place of work and in more recent times there have been significant changes to both software and hardware, resulting in increased sensitivity and specificity. It is these significant changes and advances in the sensitivity and specificity that I believe Morris (2010) was alluding to, in part, when criticising the way in which the COMICE trial selected its MRI patient data. In trying to incorporate such a large sample the COMICE study appears to have been too broad in both the time over which data was evaluated and in how studies were compared (multi-centricity and equipment).

In order to make direct comparisons or determine validity between studies it is first necessary to establish consistency or what is standard practice. At present current MRI breast practice typically involves the use of dynamic contrast-enhanced (DCE) sequences. Current breast MRI practice in New Zealand are those recommended by the New Zealand guidelines group of behalf of the Ministry of Health (2009) (NZGG, 2009). These guidelines incorporate and closely emulate the breast imaging MRI guidelines used in other countries such as those set by the American College of radiologists (ACR) (Lee, 2010), The American Cancer Society (ACS) (Saslow, 2007) and European Society of Breast Imaging (Mann, 2008). The main consensus between these protocols is to use the ACR reporting format, BI-RADS MRI Lexicon.

The American and European breast MR guidelines are well established and available for any imaging centre to use. From observing the literature these guidelines were clearly considered the status quo until the more recently published and previously mentioned British COMICE trial (Turnbull, 2010) provoked its controversy questioning the role that MRI has to play when a patient has a diagnosis of breast cancer. The COMICE trials criticism of the significance of detecting smaller less invasive cancers that would otherwise be missed



appears to have undermined both of these guidelines. While its influence on re-excision rates was comprehensive in its size and clear in its findings this study did not take into account long term survival rates. Clearly an idea for future research could be to evaluate women pre and post-operatively at given time intervals (Five year survival), in determining the effectiveness of neo-adjuvant treatment when MRI has identified additional pre-operative foci.

When imaging patients with DCE sequences some form of consistency is necessary in order to draw comparisons with any research being undertaken. Again in assessing clinical relevance of research it is first necessary to find similarities or consistency between the research being reviewed and what is achievable in practice. In New Zealand the MRI imaging standard for clinical scanning involves the use of 1.5 Telsa or in some cases 3 Telsa scanners. MRI breast imaging should ideally be performed using dedicated breast imaging coils, scanning sequences allowing good spatial and temporal resolution and where possible reported with the use of computer aided diagnosis (CAD). Radiologist experience is a crucial factor that can influence the accuracy of breast MRI and consistency can be achieved in part by using a standardised reporting format such as the ACR's BI-RADS MRI Lexicon. All imaging undertaken in this research was reviewed by one Radiologist with accuracy reviewed by way of pathology verification, audit and peer review.

CAD systems are a newer technology in breast MRI and not included in any of the guidelines or literature reviewed regarding the sensitivity and specificity research of breast MRI in combination with DWI and/or spectroscopy. The incorporation of CAD in assessing breast MRI requires a substantial financial outlay involving both specialised computer hard and software in what are standalone workstations. In literature that has assessed the role of CAD in clinical practice Dorrius et al. (2011) described how it can increase the sensitivity further when an experienced radiologist is not available. Williams et al. (2007) found significant increases in breast lesion specificity, but the key message in the literature surrounding CAD is that it brings more consistency and as such could be especially beneficial when assessing both pre and post-operative breast MRI.

In the research undertaken with this study DCE imaging was undertaken with a 1.5 Tesla MR scanner (GE Signa HDx) and using a high resolution dedicated 8 channel breast coil. Both breasts were scanned simultaneously using a three-dimensional gradient echo sequence (VIBRANT™, GE healthcare) acquiring a high resolution (spatial) data set with robust fat

saturation in a time that gave good temporal resolution. All reporting was undertaken by an experienced radiologist with more than ten years MRI experience using a CADstream™ version 4.1 workstation (GE healthcare). This protocol meets and exceeds the minimum standards for breast MRI imaging as set by the New Zealand guideline group (2009).

### *Earlier detection or changing the way breast Cancer is detected*

There are two clear positions as to the relevance of detecting small non-invasive breast cancers. The COMICE trial makes little of additional small cancers with their findings stating that MRI with its superior sensitivity does not affect re-excision rates. This point of view is not without faults due to the arguments about long-term follow up.

The position to treat each area of cancer as having the potential to become invasive is covered in the literature with regards to identifying DCIS is described by Mossa-Bassa et al. (2010), Kuhl et al. (2007) and Boyd (2011). These authors have all highlighted the role that MRI has in identifying DCIS, further describing the significance that non-invasive DCIS has in becoming invasive disease.

The main point of difference between these two positions is whether these cancers are detected during screening with MRI and the only findings or if they are additional lesions identified with preoperative MRI when a larger more invasive focus has been identified. When these lesions are additional the findings of the COMICE trial were that the identification of these lesions does not affect re-excision rates by the way of positive margins seen in post-operative pathology samples. DCIS lesions can be widely distributed within a breast and not necessarily adjacent to these larger lesions, which are commonly removed by techniques that are deemed to be conservative like lumpectomies or wide local excisions.

What is the consequence of discovering smaller adjacent non-invasive cancers? Neither position appears to be incorrect with the treatment of larger more invasive cancers a priority, which is typically followed by neo-adjunct therapy that may negate the threat posed by any adjacent smaller cancers left behind. These patients are typically followed up as part of their treatment and as such could be evaluated for the long term effects that Kuhl et al. suggest was lacking from the COMICE trial.

Because of its high sensitivity breast MRI is detecting smaller cancers than ever before. With their detection the dilemma is whether, or with what frequency, these lesions need follow up and/or treatment. With the significance of these lesions thought to pose a threat to patients by

having the potential to become invasive any difficulties in lesion specificity can cloud clinical decisions requiring additional and more invasive methods of assessment such as biopsy.

### *Sensitivity and specificity*

The optimum diagnostic test requires both high sensitivity and specificity and as such can be used as tools to identify and compare programs. Again using the definitions given by of Huang et al. (2004) The sensitivity of an examination is the probability that results from imaging are positive in those patients who have the disease, which in the context of breast MRI is the number of breast lesions detected expressed as percentage of the total number of lesions known to be present within the same sample (pathology). Specificity is the probability that results from imaging are negative in patients who do not have the disease. Specificity may also be described as depicting the accuracy of an examination with regards to false positive results in a sample of patients without the condition. More specifically in breast MRI specificity relates to the ability of the technique in correctly differentiating malignant from benign breast lesions.

This study found a sensitivity and specificity of 90% and 62.5% respectively when evaluating 68 breast lesions with DCE imaging alone. In comparing these with those in the literature the sensitivity results for this sample were consistent with those quoted by Yabuuchi, et al. (2008) (88 to 100%). The specificity result of 62.5% was the same as quoted by Huang et al. (2004) in what was similar research with a marginally smaller sample size. This specificity result while consistent with Huang et al. is significantly lower than some of the other research reviewed. This comparison is consistent because of the similarities in sample (malignant vs benign and lesion type) between these two studies.

When evaluating the false negative results in this sample there is some commonality with three of the six false negatives having pathology results that classified the lesions as being ductal carcinoma in-situ (DCIS). These three false negative DCIS lesions accounted for three of the eleven or 27% of the total DCIS sample. Two of these three DCIS lesions were classified pathologically as being of an intermediate grade and clinically significant, with the literature describing these as having the potential to become invasive disease (Kuhl et al. 2007).

DCE imaging of the breast in this sample did show excellent results for invasive cancers demonstrating 100% sensitivity for all lobular lesions and 96% of those classified as being invasive ductal carcinoma (IDC).

Overall DCE breast MRI showed good sensitivity with specificity at the lower end of what is quoted in the literature. Specificity in breast MRI research appears to be highly variable in the literature (50-90%) (Pereira, et al., 2009), and in observing the different studies seems to be heavily dependent upon the exclusion criteria for the examination and mix of lesion types in the sample. This study did include a significantly higher proportion of DCIS than most of the literature quoted 11/68 (16%) (This research) vs. 4/53 (8%) (Yabuuchi, et al., 2008), with these factors more than likely accounting for and differences in specificity.

As with the sample of lesions in the research by Huang et al. the lesions in this sample were not BI-RAD's Lexicon classified and as such probably included some lesions that were BI-RAD's five and six which had the potential to include some sample bias. This sample bias has affected the confidence levels in this research by influencing the number of benign lesions that were used to calculate specificity. In order to make a more direct comparison with more of the reviewed literature an ideal sample or an area for a future improvement would be to include only BI-RAD's class three and four lesions.

While there were similarities in sample type, a significant point of difference between this and the study by Huang et al. was that this research was undertaken using CADstream, which doesn't appear to have influenced the specificity as much as Williams et al, (2007) have suggested with their research. This is strictly an observation as it is difficult to quantify without direct comparison between CAD and non-CAD reporting method, which was outside the scope of this particular research.

#### *Diffusion weighted imaging (DWI)*

When setting out to include DWI sequences in clinical breast MRI protocols the first consideration must be the technical considerations in constructing a robust sequence. Breast DWI can be used in two ways, by allowing assessment of lesion morphology and/or quantitatively by applying ADC thresholds to differentiate whether a lesion is malignant or benign.

In constructing a robust DWI sequence the first decision that must be made is which, b-value pairing to use. Breast DWI can be performed using two or more b-values with parameters optimised to eliminate image artefacts with maximum spatial and temporal resolution.

#### *b-value selection*

When performing DWI sequences, image weighting is controlled by a factor known as the b-value. The selection of the b-value effectively sensitises the sequence to cellular water by applying a strong magnetic field gradient pulse. The selection of b-value is an important feature in the design of a DWI sequence. Lower b-values allow higher lesion conspicuity but are more susceptible to artefacts, while selection of higher b-values decreases artefacts allowing more accurate ADC characterisation.

Research that was very similar to that undertaken is that by Pereira et al. (2009). This research involved a similar patient sample and was undertaken on the same type of scanner (GE Signa), which made the technical parameters of the DWI sequence used easier to replicate clinically. This study was comprehensive in its analysis of b-values and examined the use of multiple values versus pairings before recommending the pairing of 0 and 750s/mm<sup>2</sup> describing how when used in conjunction with an ADC threshold of  $1.24 \times 10^{-3}$  mm<sup>2</sup>/s were able to achieve sensitivity and specificity of 92% and 96% respectively.

Based upon the results of the research by Pereira, et al. (2009) this b-value pairing (0 and 750s/mm<sup>2</sup>) was used with the intention of evaluating and comparing these sensitivity and specificity values clinically. There is a large amount of literature including Yabuuchi, et al (2008) that advises the use of at least three other b-values, using a mean ADC value. While it is possible to run these three DWI sequences it makes the examination more difficult both clinically and impacting patients due to the extra time penalties involved with running extra sequences and additional post processing, and as Pereira et al. (2009) have demonstrated specificity does not appear to change.

Once b-value choice was finalised the other important factors involved in designing a robust sequence were repetition time (TR) selection and matrix size including which directions to phase the sequence. After some trial and error and with some application specialist input the matrix was reduced in size from 128<sup>2</sup> to 128 (phase) x 92 (frequency), which drastically reduced any distortion experienced in the frequency plane. Different TR's were tried with 2800ms preferred. This proportionally low TR was chosen to reduce T2\* artefacts.

### *DWI Sensitivity and specificity*

In reviewing the literature there appears to be two approaches to sample selection in order to determine specificity and PPV for breast MRI. One approach is to include only lesions that have been pathology verified and the other being to include lesions that appear to be benign with other imaging modalities. The lesions examined in this sample were not formally reported using BI-RADS format and as it involved retrospective analysis only included pathology confirmed lesions. Reported differences in specificity of breast MRI in the literature does appear to be influenced by both the number of benign lesions and the types of malignant lesions being sampled. Some research, such as that by Partridge, et al. (2009) have samples of lesions that exclude biopsy proven lesions (BI-RADS 6) this appears to be in order to control for bias but as it was a retrospective study more likely used to attain a better ratio of benign and malignancy for specificity calculations.

### *DWI as a standalone procedure*

DWI in this examination was assessed retrospectively following positional correlation with DCE and the associated pathology results. Lesion morphology was not necessarily required on the DWI images but focused upon identifying lesions by their ADC thresholds. DCE images were assessed on a CADstream version 4.1 workstation. The next generation version 5.4 CADstream software allows more accurate lesion correlation with ADC values being available when hovering over a suspected lesion on the DCE subtraction images (MERGE Healthcare, Chicago, Illinois, USA).

The results undertaken in this research with a sample of 68 lesions, DWI imaging alone demonstrated a sensitivity and specificity of 98% (95%CI: 90-100%) and 63% (95%CI: 26-89%) respectively. The sensitivity of 98% achieved for DWI alone was 8% higher than the sample with DCE imaging and consistent with the literature. In similar but not identical research Partridge et al. (2010) achieved a sensitivity of 96% and specificity of 56% when using a b-value pairing of 0 and 600 s/mm<sup>2</sup>. The specificity achieved in this research of 63% demonstrated almost identical specificity to the sample evaluated with DCE alone and an increase of approximately 7% over the sample evaluated by Partridge et al. (2010).

The DWI sample included two false negatives, both of which were lobular lesions. These two lesions were just outside the thresholds used for classifying malignancy. Had these two false positives been correctly classified by DWI the specificity would have been in the same order

as the research by Guo et al. (2002). Increased specificity in this sample could have been achieved by re-evaluating the ADC thresholds. Adjustment of the ADC thresholds was outside of the scope of this research and with any adjustment in ADC tradeoffs would be potentially increasing the number of false positives by causing increased ADC threshold overlap.

The three false positives in the DWI sample included lesions with ductal hyperplasia, chronic inflammatory and fibrotic changes. The chronic inflammatory false positive DWI lesion was also one of the three false positive DCE lesions, while the ductal hyperplasia and fibrotic lesions were correctly diagnosed using DCE breast MRI.

DWI alone showed 100% sensitivity to all IDC and DCIS lesions in the sample. In detecting 100% of the DCIS sample DWI increases the overall sensitivity of the examination. The significance of detecting DCIS has been debated in the literature with Turnbull (2010) in the COMICE study stating that detecting smaller less invasive cancers does not appear to influence re-excision rates which, Kuhl et al. (2007) state have the potential to become invasive disease.

#### *DCE and DWI imaging combined*

The intention of this research was to examine the use of DWI not as a standalone procedure but in order to complement DCE imaging which, as the literature demonstrates, has very good sensitivity. It was the hypothesis of this research that the specificity of the DWI sequence and ADC quantification that would be the most important factor in recommending its inclusion into routine protocols for diagnosis of breast cancer.

When evaluating the combination of the DWI and DCE samples, two rules were made. Rule a) had perfect sensitivity with poor specificity (38% and poor 95% confidence levels. Rule b) was more useful with Sensitivity and specificity of 90% (95% CI: 79-96%) and 86% (95% CI: 47-99%) respectively, compared to that by Pereira, et al. (2009) with 92.3% sensitivity (95% CI 75.9-97.9%) and 96.2% specificity (95% CI 91.1-99.3%). While these results were different they were within the confidence intervals of this research (sensitivity: 79-96%, specificity 47-99%). Pereira et al. achieved closer 95% confidence intervals due to a better ratio of malignant to benign lesions with this being a consequence of different methods of sample selection between this research and that undertaken as part of this research. In the

sample of DCE and DWI imaging combined there was one lesion that tested false positive on both sequences with this being classified as being chronic inflammatory change.

### *Pathology specific DWI*

One of the questions that this research was seeking to clarify was whether or not quantitative DWI was able to provide information enabling the differentiation of breast lesions.

The results of this research found good separation of benign and malignant thresholds but demonstrated no significant separation of thresholds for any of the four malignant types of lesions in this sample. These threshold separations are evident in the threshold comparison graph. These findings were consistent with those by Park et al. (2007) who also disregarded any association between lesion types in a similar study evaluating quantitative DWI.

In research that did find an association between histological types, Woodhams et al. (2009) found significant differences in ADC thresholds between mucinous and other malignant lesions. These differences showed significant crossover between benign and malignancy which if the case would influence confidence in this technique where a significant number of these lesions were contained in the sample. This research contained a significantly smaller sample of mucinous carcinoma (n=4) but differed in its findings with all of these lesions conforming to the threshold separation demonstrated by the other benign and malignant lesions contained within this sample. A point of difference between these studies was that Woodhams et al. used only lesions classified as being pure Mucinous lesions (n=15), while the sample used in this research contained both pure lesions (n=2) with the remaining 2 lesions in this sample being mixed mucinous.

Razek et al. (2010) in their research found differences in ADC values depending upon nuclear grading between lesions of the same pathology (IDC). This separation by nuclear grading was unfortunately outside the scope of this research, but could provide an area for future evaluation for research especially between non-invasive lesions containing DCIS.



### *Gadolinium*

Gadolinium has been found to reduce perfusion artefacts and reduce the mean ADC threshold for breast lesions on DWI sequences (Peters, Vincken, van den Bosch, Luijten, Mali, & Bartels, 2010). The significance of these factors is that any ADC thresholds used must come from research involving similar b-values and also post administration of Gadolinium in order to reproduce both sensitivity and specificity. The thresholds in this research were from Pereira, et al. (2009) and thus controlled for both perfusion and ADC threshold reductions.

### *High field DWI*

Results from high field DWI is mixed with Peters, Vincken, van den Bosch, Luijten, Mali, & Bartels (2010) suggesting that DWI at 3 Tesla is no better than at 1.5 Tesla because of technical problems (fat saturation and shimming) canceling out any improvements in SNR. Research with results contradictory to Peters et al (2010) was that by Lo et al. (2009) who stated that 3 Tesla increases in SNR gave significant increases in sensitivity and specificity of 90 and 91% respectively further stating that because of the increased SNR greater b-values were able to be used without perfusion penalties.

### *Spectroscopy*

Breast spectroscopy is a sequence that requires lesion visualisation in three imaging planes and is used primarily for lesion differentiation. Lesion differentiation in breast spectroscopy is possible because of the sequences ability to interrogate a lesions metabolic structure or more specifically if choline is present within the sample.

The themes presenting in the literature surrounding breast spectroscopy include the pathological process by which choline is produced in cancerous lesions, the position and composition of the choline peak on the MR spectrum as well as factors affecting the sensitivity and specificity of this sequence such as lipid contamination of the voxels and whether breast spectroscopy is suitable for use on mass like and non-mass like lesions.

Literature involving high-field breast MR spectroscopy was identified and while outside the scope of this research appears to be of significance to the future of this sequence.

Choline metabolites are not normally detectable in breast tissues at 1.5 Tesla and when present in the MR spectra indicate malignancy as a by-product of cellular membrane breakdown.

### *Spectroscopy sensitivity and specificity*

The sensitivity of breast spectroscopy to non-mass like lesions is potentially very poor as typical sequences involve the use of large voxels in the region of 1 to 2 cm<sup>3</sup> in order to fulfil SNR requirements in clinically acceptable times. There is research supporting the evaluation of non-mass like lesions with spectroscopy such as that by Bartella et al. (2007), although this is contradicted by others including Tozaki and Fukama (2009) who stated that breast spectroscopy performed poorly for non-mass like breast lesions. This along with the practical disadvantages of undertaking non-mass like spectroscopy led to the exclusion of these from the sample used for this research. Breast spectroscopy of mass like lesions shows more promise clinically with single voxel spectroscopy at 1.5 Tesla able to accurately differentiate lesions in the order of 1.5cm<sup>3</sup>.

The spectroscopy reviewed as part of this research involved 22 patients all of whom had mass like lesions. As all of the lesions sampled were over 1cm<sup>3</sup> there was a bias towards malignancy with sensitivity being 100%. Because of this predominance of malignant over benign lesions the population of this sample was too small to make a valid estimate of specificity. The confidence levels for PPV (90%, 95%CI: 68-98%) were more significant than those achieved for specificity.

The sensitivity and PPV achieved in this research were very similar to that achieved by Bartella et al (2004) (100% sensitivity and 91% PPV). The one false positive in this spectroscopy sample was a significantly sized lesion that was pathologically diagnosed as being an area of benign fibrotic change. This lesion gave a strong but wide choline peak at approximately 3.26ppm and was a lesion that also tested positive using DCE breast MRI. The DCE results for this sample were the same and as such have exactly the same sensitivity, specificity and PPV. The DWI results were marginally different with the false positive spectroscopy result being correctly diagnosed in this sample. One Lobular lesion in this sample was incorrectly diagnosed as being negative with an ADC threshold of 0.645 when the lower limit for malignancy is  $0.68 \times 10^{-3} \text{mm}^2/\text{s}$ .

Identified in the literature was the issue of an additional secondary choline peak identified at 3.28ppm which is associated with benign breast tissue (Tozaki and Fukama, 2009). This secondary peak is described as being significantly smaller than that centred at 3.26ppm and as such does not pose a significant problem using single voxel techniques to assess breast lesions at 1.5 Tesla.

## Limitations:

### *Context bias*

When examining any potential weaknesses in design the approach of this research in using only pathology confirmed lesions has limited the number of benign lesions within this sample. Yabuuchi, et al. (2008) discussed these limitations in their research as context bias potentially inflating sensitivity at the expense of specificity. The lack of pathology confirmed benign lesions was more evident with the spectroscopy sample resulting in the study being unable to give specificity results, because of a larger than optimal margin between the 95% confidence intervals of that particular sample.

One approach in the literature to improve confidence levels is to adopt the BI-RADs format excluding any lesions identified previously as being pathology confirmed malignancies BI-RADs 6. Other authors advocate the use of benign lesions that have not undergone pathology confirmation if they have been followed up over a period of time without demonstrating change.

### *External Validity*

As this research was undertaken at a single centre only, this reduces the external validity or reproducibility of these results at different centres where different equipment or radiologist experience may be a factor in outcomes. The focus or context of this study was always intended to be for a single centre only as a program or protocol evaluation.

Radiologist experience however does not necessarily equate with quality without controls such as peer review and audit in order to prevent mistakes being made. Radiologist experience was not of particular concern for this research due to the significant breast MRI experience of the radiologist involved and as highlighted in the literature controlled for by the use of computer-aided diagnosis.

## Conclusion

Current breast MRI practice in New Zealand closely emulates the American and European guidelines. The most commonly used protocols used for imaging breasts using these guidelines are dynamic contrast enhanced (DCE) sequences that involve the intravenous injection of contrast and generation of contrast kinetic graphs to differentiate any suspicious lesions. Breast MRI has been widely researched and is described as having high sensitivity to breast lesions. Specificity of breast MRI (DCE) is more debatable with wider variations reported resulting in false positives.

This research set out to examine the hypothesis that the introduction of two additional breast MRI sequences, DWI and spectroscopy to routine protocols in order to evaluate their effect on sensitivity and specificity. This research was undertaken on a 1.5 Tesla MRI scanner using only pathology confirmed lesions. This selection of lesions included a wide range of all of the commonly presenting malignant lesions (Lobular, Invasive ductal, DCIS and Mucinous).

In order to fully evaluate the performance of DWI and spectroscopy it was firstly necessary to determine the sensitivity and specificity of DCE MRI, which gave sensitivity and specificity of 90% and 62.5% respectively. The sensitivity results were consistent with much of the literature and using these baseline results for DCE MRI, further reinforces the argument of high sensitivity vs. lower specificity. While the specificity achieved for this research was at the lower end of that reported it was consistent with the one other article (Huang et al. 2004) that used only pathology confirmed lesions. The lower specificity results of this research were also more than likely affected due to the inclusion of a significantly larger proportion of DCIS lesions in comparison to other similar studies. When evaluating the same patient sample with DWI, sensitivity was increased further to 98% with specificity being 63%. The same sample evaluated for both DWI and DCE imaging gave both sensitivity and specificity of 88% with only one lesion giving a false positive result.

The variations in sensitivity and more particularly specificity in the literature appear to be heavily dependent upon method and lesion type. DWI and spectroscopy were initially introduced to routine breast MRI protocols based upon reviews of the literature, and while most researchers attempt to control for external validity there are so many factors involved (research design, equipment, lesion types and reviewer experience) that no two articles appear to give the same results. This thesis was undertaken to give context to the prior

research and identify whether the inclusion of these two sequences does result in increasing the specificity thus validating the continued inclusion of these imaging sequences in a clinical setting. MRI is a rapidly evolving imaging modality and using the results of this research the argument for the continued inclusion of DWI sequences is a strong one due to significant increases in specificity. The increased specificity of DWI to DCIS was a significant finding and because of this the greatest factor in the increases between this and the DCE sample. This increase in specificity with DWI and DCE sequences combined signifies what appears to be a significant step forward in the accuracy of breast MRI.

The spectroscopy sample used was significantly smaller than that used for the DCE and DWI samples with only mass-like lesions ( $< 1\text{cm}$ ) evaluated. The sensitivity of this sample was 100% respectively. Specificity calculations for this population were unreliable being significantly affected by context bias with only 3 of the 22 lesions sampled being benign. Positive predictive value for this sample was particularly high being 90%. Spectroscopy in the form it was evaluated is limited to non-mass like lesions ( $< 1\text{cm}$ ) and as such did not include any smaller lesions (DCIS) that would prove more difficult to diagnose without biopsy.

The high positive predictive value of spectroscopy could prove beneficial to patients as an alternative to a more invasive biopsy but this would need to be evaluated in a prospective outcome study.

## Recommendations

DWI using a b-value pairing of 0 and 750 s/mm<sup>2</sup> increases both the sensitivity and specificity of breast MRI to malignancy in a clinically manageable time of less than four minutes and as such warrants its continued inclusion to routine breast MRI protocols.

Spectroscopy demonstrates both high sensitivity and PPV to breast lesions but with the practical considerations associated with imaging times of between five to eight minutes needs further evaluation before being undertaken routinely.

The inclusion of DWI in combination with routine breast MRI protocols has resulted in an increase in specificity of 23%, which long-term has the potential to both decrease the need for invasive breast biopsies while further increasing the sensitivity of MRI to breast lesions.

Areas for future research into breast MRI remain and these include potential differences in ADC thresholds for lesions with different nuclear grading and whether HER2 status has any association with spectroscopy results.

In order to better evaluate the true specificity of breast spectroscopy a study incorporating a larger sample of benign lesions is required to improve confidence levels.

## References

- Anderson, G. (1998). *Fundamentals of educational research*. London: Falmer Press.
- Bartella L, T. S. (2007.). Enhancing nonmass lesions in the breast: evaluation with proton (1H) MR spectroscopy. *Radiology* , 245(1), 80-87.
- Bartella, L. M. (2004). Proton MR spectroscopy with Choline peak as malignancy marker improves positive predictive value for breast cancer diagnosis. *Radiology* , 239 (3), 686-692.
- Beatty, D. P. (2007). Contrast-enhanced breast magnetic resonance imaging: the surgical perspective. *The American journal of surgery* , 193, 600-605.
- Bolan, P. M. (2003). In vivo quantification of Choline compounds in the breast with 1H MR spectroscopy. *Magnetic resonance in medicine* , 50, 1134-1143.
- Bolan, P. N. (2005). Imaging in breast cancer: Magnetic Resonance Spectroscopy. *Breast Cancer Research* , 7, 149-152.
- Boyd, S. (2011). Evaluation of DCIS with MRI . *Educational lecture given at Dunedin Hospital*. Dunedin.
- Bushong, S. C. (2003). *Magnetic Resonance Imaging* (3rd Edition ed.). St Louis, Missouri, United States of America: Mosby Inc.
- Campbell, I. (2010). Comparative effectiveness of MRI in breast cancer (COMICE trial: a randomised controlled trial (independent commentary). *Breast cancer research review* (3), 3.
- Carbonaro, L. P. (2011). Breast MRI using a high relaxivity contrast agent. *American journal of roentgenology* , 196, 942-955.
- Cohen, L. M. (2000). *Research methods in education* (5th ed.). New York: Routledge Farmer.
- Davidson, E. H. (2007). *Surveillance of women at high risk of breast cancer*. Christchurch school of medicine and health sciences, Department of public health and general practice. Christchurch: New Zealand Health Technology Assessment (NZHTA).
- Dorrius, M. v. (2011). Computer-aided detection in breast MRI: a systemic review and meta-analysis. *European radiology* , 21 (8), 1600-1608.
- Foundation, N. Z. (2010, May 16). Retrieved May 16, 2010, from New Zealand Breast Cancer Foundation Web site: <http://www.nzbcf.org.nz/>
- Geraghty, P. V. (2008). MRI and 1H MRS of the breast. *Breast* , 14, 574-580.

- Gruber, S. D. (2011). Three-dimensional proton MR spectroscopic imaging at 3T for the differentiation of benign and malignant breast lesions. *Radiology* , 261 (3), 752-761.
- Guo, Y. C. (2002). Differentiation of clinically benign and malignant breast lesions using diffusion-weighted imaging. *Journal of magnetic resonance imaging* , 16, 172-178.
- Haddadin, I. M. (2009). Metabolite quantification and high-field MRS in breast cancer. *NMR Biomed* , 22 (1), 65-76.
- Hall, F. (2010). Identification, biopsy and treatment of poorly understood premalignant, in-situ, and indolent low-grade cancers. *Radiology* , 254, 655-659.
- Hart, C. (2006). *Doing your masters dissertation*. London: Sage Publications.
- Huang, W., Fisher, P. R., Tudorica, L. A., O'Hea, B., & Button, T. M. (2004.). Detection of breast malignancy: Diagnostic MR protocol for improved specificity. *Radiology* . , 232, 585-591.
- Jones, D. H. (1999). Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine* , 42, 515-525.
- Katz-Brull, R., Lavin, P.T., Lenkinski, R. F. (2002) Clinical utility of proton magnetic resonance spectroscopy in characterizing breast lesions. *Journal of the National Cancer Institute*, 94 (16), 1197-1203.
- Klomp, D. v. (2010). 31P MRSI and 1H MRS at 7T: initial results in human breast cancer. *NMR in Biomedicine* , 24, 1337-1342.
- Kuhl, C. S. (2007). MRI for diagnosis of pure ductal carcinoma in situ. *Lancet* , 370, 485-492.
- Kuroki, Y. N. (2008). Advances in breast MRI: Diffusion-weighted imaging of the breast. *Breast cancer* , 15, 212-217.
- Lalkhen, A.G., & McCluskey, A.M. (2008). Clinical tests: Sensitivity and Specificity. *Journal of continuing education in anaesthesia, critical care and pain* , 8(6), 221-223.
- Lee, C. D. (2010). Breast cancer screening with imaging: Recommendations of the society of breast imaging and the ACR on the use of mammography, breast MRI, Breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *Journal of the American college of radiology* , 7 (1), 18-27.
- Lo, G. A. (2009). Diffusion-weighted magnetic resonance imaging of breast lesions: First experiences at 3T. *Journal of computer assisted tomography* , 33 (1), 63-69.
- Mann, R. K. (2008). Breast MRI: Guidelines from the european society of breast imaging. *European Radiology* , 18 (7), 1307-1318.



- Marini, C., Iacconi, C., Giannelli, M., Cillotti, A., Moretti, M., & Bartolozzi, C. (2007.). Quantitative diffusion-weighted MR imaging in the differential diagnosis of breast lesion. *European Radiology* , 17, 2646-2655.
- MedCalc. (2013, 06, 27). Retrieved from Easy to use statistical software: <http://www.medcalc.org/?gclid=CKft1qe3grICFSZxQgodxhAAjw>
- Morris, E. (2010). Should we dispense with preoperative breast MRI? *The Lancet* , 375 (9714), 528-530.
- Mossa-Bassa, M. F. (2010). Ductal carcinoma in situ of the breast: MR imaging findings with histopathologic correlation. *Radiographics* , 30, 1673-1687.
- Neuman, W. (2003). *Social research methods. Qualitative and quantitative approaches*. USA: Allyn & Bacon.
- NZGG. (2009). *Management of early breast cancer*. Wellington: New Zealand Ministry of Health.
- Obuchowski, N. (2003) Receiver operating characteristic curves and their use in radiology. *Radiology*, 229, 3-9.
- Ogura, A. H. (2008). The effect of susceptibility of gadolinium contrast media on diffusion-weighted imaging and the apparent diffusion coefficient. *Academic radiology* , 15 (7), 867-872.
- Park, M. C. (2007). The role of diffusion-weighted imaging and the apparent diffusion coefficient (ADC) values for breast tumors. *Korean journal of radiology* , 8 (5), 390-396.
- Partridge, S. D. (2009). Quantitative Diffusion-Weighted imaging as an adjunct to conventional breast MRI for improved positive predictive value. *American journal of Roentgenology* , 193, 1716-1722.
- Partridge, S. M. (2010). Apparent diffusion coefficient values for discriminating benign and malignant breast MRI lesions: Effects of lesion type and size. *American Journal of Roentgenology* , 194, 1664-1673.
- Partridge, S., DeMartini, W., Kurland, B., Eby, P., White, S., & Lehman, C. (2010). Differential Diagnosis of Mammographically and Clinically Occult Breast Lesions on Diffusion-Weighted MRI. *Journal of Magnetic Resonance Imaging* , 31, 562-570.
- Peat, J. (2001). *Health science research*. Sydney: Allen & Unwin. Sydney: Allen & Unwin.
- Pereira, F., Martins, G., Figueredo, E., Domingues, M., Domingues, R., Barbosa da Foneseca, L., et al. (2009.). Assessment of breast lesions with diffusion-weighted MRI: Comparing the use of different b values. *American Journal of Roentgenology* . , 193, 1030-1035.

- Peters, N., Vincken, K., van den Bosch, M., Luijten, P., Mali, W., & Bartels, L. (2010). Quantitative Diffusion Weighted Imaging for Differentiation of Benign and Malignant Breast Lesions: The Influence of the Choice of b-values. *Journal of Magnetic Resonance Imaging* , (31), 1100-1105.
- Raza, S. B. (2010). *Specialty imaging Breast MRI a comprehensive imaging guide*. Salt Lake City: Amirsys.
- Razek, A. G. (2010). Invasive ductal carcinoma: correlation of apparent diffusion coefficient value with pathological prognostic factors. *NMR in biomedicine* .
- Rijnsburger, A. O.-L. (2010). BRCA1-Associated breast cancers present differently from BRCA2-Associated and familial cases. *Journal of clinical oncology* , 28 (36), 5265-5273.
- Roebuck, J. C. (1998). Human breast lesions: Characterisation with proton MR spectroscopy. *Radiology*, 209(1), 269-275.
- Rubesova, E. G. (2006). Quantitative diffusion imaging in breast cancer. *Journal of Magnetic resonance imaging* , 24, 319-324.
- Sardenelli, F. F. (2009). In vivo proton breast MR spectroscopy of the breast using the total Choline peak integral as a marker of malignancy. *American journal of roentgenology* , 192, 1608-1617.
- Saslow, D. B. (2007). American cancer society guidelines for breast screening with MRI as an adjunct to mammography. *CA: A cancer journal for clinicians* , 57, 75-89.
- Stanwell, P., & Mountford, C. (2007). In Vivo Proton MR Spectroscopy of the Breast. *RadioGraphics* , 27 (Special Issue), S253-S267.
- Su, M. (2008). Detection of choline signal in human breast lesions with chemical shift imaging. *Journal of magnetic resonance imaging* , 27 (5), 1114-1121.
- Suzuki, S., K. Y. (2007). Detecting breast cancer with non contrast MR imaging: Combining diffusion-weighted and STIR imaging. *Magnetic resonance in medical sciences* , 6 (1), 21-27.
- Tozaki, M. H. (2010). <sup>1</sup>H MR spectroscopy of invasive ductal carcinoma: Correlations with FDG PET and histologic prognostic factors. *American journal of roentgenology* , 194, 1384-1390.
- Tozaki, M., & Fukuma, E. (2009.). <sup>1</sup>H MR spectroscopy and diffusion-weighted imaging of the breast: Are they useful tools for characterizing breast lesions before biopsy? *American Journal of Roentgenology*. , 193, 840-849.
- Tozaki, M. M. (2010). Current status and future prospects of proton MR spectroscopy of the breast with a 1.5T MR unit. *Journal of oncology*, 1-10.

Trochim, W. (2006, October 20). *Research Methods Knowledge Base*. Retrieved May 25, 2010, from Research Methods Knowledge Base:  
<http://www.socialresearchmethods.net/kb/intreval.php>

Tse, G. Y. (2007). In vivo proton magnetic resonance spectroscopy of breast lesions. *Breast cancer response to treatment* , 104, 249-255.

Tsushima, Y. T.-T. (2009). Magnetic resonance (MR) differential diagnosis of breast tumors using apparent diffusion coefficient (ADC) on 1.5-T. *Journal of magnetic resonance imaging* , 30, 249-255.

Turnbull, L. B. (2010). Comparative effectiveness of MRI in breast cancer (COMICE) randomised controlled trial. *The Lancet* , 375 (9714), 563-571.

Vassarstats. (2012, 5 14). Retrieved from Clinical calculator 1:  
<http://vassarstats.net/clin1.html>

Wang, J. T. (2008). Proton MR spectroscopy in normal breasts between pre- and post menopausal women. *American ournal of roentgenology* , 190 (2), 505-510.

Williams, T. D. (2007). Breast MR imaging: Computer=aided evaluation program for discriminating benign from malignant lesions. *Radiology* , 244 (1), 94-103.

Woodhams, R. K. (2009). Diffusion-weighted imaging of Mucinous carcinoma of breast. *American journal of Roentgenology* , 193, 260-266.

Woodhams, R. K. (2009). Identification of residual breast carcinoma following neoadjuvant chemotherapy. *Radiology* , 254 (2), 357-366.

Woodhams, R. M. (2005). Diffusion-weighted imaging of malignant breast tumors. *Computer assisted tomography* , 29 (5), 644-649.

Woodhams, R. R. (2011). Diffusion-weighted imaging of the breast: Principles of clinical applications. *Radiology*, 31, 1059-1084.

Yabuuchi, H., Matsuo, Y., Okafuji, T., Kamaitani, T., Soeda, H., Setoguchi, T., et al. (2008). Enhanced mass on contrast-enhanced breast MR imaging: Lesion characterization using combination of dynamic contrast-enhanced and diffusion-weighted MR images. *Journal of Magnetic Resonance Imaging* . , (28), 1157-1165.

Yoo, J. W. (2010). Can MR imaging contribute in characterising well-circumsised breast carcinomas? *Radiographics* , 30, 1689-1704.

## Appendix

### Sensitivity and specificity Tables

DCE:

Values entered:

DCE	Breast lesion		Totals
	Benign	Malignancy	
Test Positive	3	54	57
Test Negative	5	6	11
Totals	8	60	68

DWI:

Values entered:

DWI	Breast lesion		Totals
	Benign	Malignancy	
Test Positive	3	59	62
Test Negative	5	1	6
Totals	8	60	68

DCE and DWI:

Values entered:

DCE & DWI (Rule b)	Breast lesion		Totals
	Benign	Malignancy	
Test Positive	1	54	55
Test Negative	7	6	13
Totals	8	60	68

Spectroscopy:

Values entered:

Spectroscopy	Breast lesion		Totals
	Benign	Malignancy	
Test Positive	1	18	19
Test Negative	2	0	2
Totals	3	18	21

ROC curve tables:

### Empirical ROC DWI

Variable	ADC
Classification variable	Histology

Sample size		68
Positive group :	Histology = 1	60
Negative group :	Histology = 0	8

Disease prevalence (%)	unknown
------------------------	---------

### Area under the ROC curve (AUC)

Area under the ROC curve (AUC)	0.844
Standard Error <sup>a</sup>	0.0892
95% Confidence interval <sup>b</sup>	0.735 to 0.920
z statistic	3.854
Significance level P (Area=0.5)	0.0001

<sup>a</sup> DeLong et al., 1988

<sup>b</sup> Binomial exact

### Youden index

Youden index J	0.6250
Associated criterion	≤1.3

### Criterion values and coordinates of the ROC curve [\[Hide\]](#)

Criterion	Sensitivity	95% CI	Specificity	95% CI	+LR	-LR
<0.645	0.00	0.0 - 6.0	100.00	63.1 - 100.0		1.00
≤0.9	31.67	20.3 - 45.0	100.00	63.1 - 100.0		0.68
≤0.93	35.00	23.1 - 48.4	87.50	47.3 - 99.7	2.80	0.74
≤1	65.00	51.6 - 76.9	87.50	47.3 - 99.7	5.20	0.40
≤1.03	65.00	51.6 - 76.9	75.00	34.9 - 96.8	2.60	0.47
≤1.1	75.00	62.1 - 85.3	75.00	34.9 - 96.8	3.00	0.33
≤1.11	78.33	65.8 - 87.9	62.50	24.5 - 91.5	2.09	0.35
≤1.3	100.00	94.0 - 100.0	62.50	24.5 - 91.5	2.67	0.00
≤1.6	100.00	94.0 - 100.0	0.00	0.0 - 36.9		

(MedCalc, 17/06/2013)