



IMAGE TEXTURE ANALYSIS OF TRANSVAGINAL ULTRASOUND IN THE DIAGNOSIS OF OVARIAN LESIONS

Institute of Medical Engineering and Medical Physics, Cardiff School of Engineering, Cardiff University.

By: Rana Hussain ALdahlawi

February 2016

A thesis submitted to the Cardiff University in candidature for the degree of Doctor of Philosophy

Dedication

I proudly dedicate my dissertation to my beloved parents, Hussein and Fawziyah, who are the soul of my life. Their words of encouragement and constant inspiration to better myself and aim higher ring in my ears. Thanks also to my sisters and brothers, who have never left my side and have always been there for me and supported me over the years.

I also dedicate it to my husband, Hatem, and my wonderful children, Lamar, Abdulmalek and Lareen, who are the joy of my life and without whom life would never be the same. They have been my inspiration in achieving my goal. Special thanks to my Nanny Michelle Thomas, who has been very helpful and caring to us throughout the years and made life easier for me.

I dedicate this work and give special thanks to my friends Sameeha, Ma'rab and Eman for being there for me throughout the entire doctorate program.

Acknowledgments

First and foremost, praise to almighty ALLAH, the beneficent and the most merciful, for giving me the opportunity and strength to go through this study. Without the help and guidance of almighty God, this work, or indeed any other work, would not have been possible to achieve.

My gratitude to Prof. Neil Pugh, who was more than generous with his expertise and precious time, and for his countless hours of guidance, reflecting, reading, encouraging, and most of all patience throughout the entire process. I will be forever influenced by his incisive ideas. This extends to Prof. Len Nokes for all his support. I could not wish for better supervisors.

I would like to acknowledge and thank the Doppler department for allowing me to conduct my research and providing any assistance requested in a professional and comfortable environment. Special thanks goes to all members of staff, namely, Mrs. Gillian Sullivin, Dr Declan Coleman, for his priceless advice on the technical aspects of texture analysis, and Dr. Paul Williams, Rhys Morris and Kate Wells for their assistance on technical aspects of ultrasound.

This appreciation is extended to Dr. Bidi Hamid, who inspired me with his exceptional work, gave me guidance to follow his work and assisted me with the technical aspects of texture analysis.

I would also like to thank Mrs. Sally Davey, the senior product specialists in Toshiba machines for her help with ASQ. I would also like to acknowledge Dr. Anju Sinha and Dr. Kenneth Lim for their assistance in recruiting the participants and in other clinical aspects of the study.

Lastly, the financial support of King Saud University in Riyadh, which provided a grant for my studies, is gratefully acknowledged.

Abstract

Ovarian cancer has the highest mortality rate of all gynaecological cancers and is the fifth most common cancer to occur in women in the UK. Amongst various imaging modalities, ultrasound is considered the main modality for ovarian cancer triage. As with other imaging modalities, the main issue is that the interpretation of ultrasound images is subjective and observer dependent. Texture analysis has been shown to have potential in the objective assessment of ovarian cancer in a preliminary study. Another form of texture analysis is Acoustic Structure Quantification (ASQ), which has been documented to have a number of successful uses in liver diseases. However, it has not been applied to ovarian lesions. Therefore, the aim of this study was to assess prospectively the diagnostic performance of texture analysis methods such as GLCM, Wavelet, and ASQ in discriminating between benign and malignant adnexal masses and between different types of benign masses and compare it to widely used scoring models.

Prior to applying ASQ to ovarian images, its reliability and repeatability were first evaluated. This includes random variation caused by the ultrasound system and the operator during image acquisition. A tissue-equivalent phantom was used in these tests. It was found that the ASQ feature demonstrated excellent repeatability for ASQ software, with all transducers showing less than 0.4% variance from the mean: thus, ASQ software is able to produce reliable ASQ output measures. When testing the factors that may influence the performance of the ASQ analysis, the results revealed that three factors do not influence the mean of the output curve: the ROI size, depth and gain setting. However, focal position has a significant effect on the mean of the output curve. Transducer frequency does not affect the output curve except when using high frequencies such as 8 MHz. Other tests were done to determine the appropriate parameters in the software to be used on images of ovarian masses.

Firstly, ASQ was applied to 45 pelvic masses. The preliminary results showed no significant difference between benign and malignant masses using the ASQ technique: therefore, the study was terminated due to failure to discriminate the benign from the malignant masses using ASQ.

Secondly, two types of textural features were investigated in this study: grey-level co-occurrence matrix (GLCM) and wavelet, as recommended by a preliminary study. A sample of 169 masses was collected from participants, of which 140 were benign and 29 were malignant by histology. In addition to texture features, other widely used scoring models were applied on the same images for comparison, namely RMI, PMI and ADNEX.

The results revealed excellent discriminatory ability in both GLCM and wavelet between malignant and cystic masses and between benign and cystic masses, with AUC of .994 and .895 for GLCM and .894 and .814 for wavelet respectively, as well as between normal and malignant tissue, with $p > .05$ and $p = .004$ in both GLCM and wavelet respectively.

Results also showed that GLCM outperformed RMI and ADNEX in distinguishing between benign and malignant masses, even when dividing the study population into pre- and postmenopausal groups. In addition, GLCM has the advantage of being objective and not operator dependent. Receiver operating characteristic (ROC) curve analysis was carried out to determine the discriminatory ability of textural features, which was found to be satisfactory.

The principal conclusion was that GLCM and wavelet features can potentially be used as computer aided diagnosis (CAD) tools to help clinicians in the diagnosis of ovarian cancer.

List of Tables.....	X
List of Figures	xiii
Abbreviations.....	xvi
1. Introduction.....	1
1.1. Research hypothesis	4
1.2. Aims and objectives	5
1.2.1. Objectives:.....	5
1.3. Thesis structure	6
2. Literature Review	8
2.1. Background	8
2.1.1. Ovarian cancer	9
2.1.1.1. Ovaries: location and appearance	9
2.1.1.2. Ovarian cancer: types	12
2.1.1.3. Ovarian Cancer: Staging.....	12
2.1.1.4. Risk Factors	15
2.1.1.4.1. Age.....	15
2.1.1.4.2. Family History.....	16
2.1.1.4.3. Height	16
2.1.1.4.4. Body Mass Index (BMI).....	17
2.1.1.4.5. Hysterectomy	18
2.1.1.4.6. Use of Hormone Replacement Therapy (HRT).....	18
2.1.1.5. Symptoms	19
2.1.2. Methods of Preoperative Evaluation.....	21
2.1.2.1. Tumour markers	22
2.1.2.1.1. Serum CA125.....	22
2.1.2.1.2. Other Biomarkers	25
2.1.2.2. Ultrasound	27
2.1.2.2.1. B-mode (brightness-mode) Ultrasound: (see figure 3)	30
2.1.2.2.2. Three-dimensional ultrasound.....	45
2.1.2.2.3. Contrast Ultrasound.....	46
2.1.2.2.4. Doppler.....	47
2.1.2.3. CT imaging.....	53
2.1.2.4. Magnetic Resonance Imaging (MRI)	54
2.1.2.5. Laparoscopy and Laparotomy	58
2.1.2.6. Texture analysis.....	59
2.1.2.6.1. Application of texture analysis to medical images.....	64
2.2. Summary	71
3. Acoustic Structure Quantification (ASQ).....	73
3.1. (A) Introduction, how ASQ works and background	73

3.1.1	How ASQ works	74
3.1.2.	ASQ background.....	81
3.2.	Repeatability and reproducibility (Phantom studies)	84
3.2.1.	Materials, methodology and statistical analysis	85
3.1.3.1.	Materials	85
3.1.3.1.1.	<i>Ultrasound scanner:</i>	85
3.1.3.1.2.	<i>Ultrasound Transducer:</i>	86
3.1.3.2.	Methods	93
3.1.3.2.1.	<i>Intra-operator Repeatability</i>	93
3.1.3.2.2.	<i>Inter-operator Reproducibility</i>	96
3.1.3.2.3.	<i>Coefficient of Variation (CoV)</i>	97
3.1.3.2.4.	<i>Bland-Altman plot</i>	97
3.1.3.3.	Results	99
3.1.3.3.1.	<i>Intra-operator repeatability</i>	99
3.1.3.3.2.	<i>Repeatability of ASQ software</i>	99
3.1.3.3.3.	<i>Inter-operator reproducibility</i>	100
3.1.3.4.	Discussion.....	103
3.3.	(B) Influence of ROI size, ROI depth, Focus, Gain setting and transducer frequency on ASQ	106
3.3.1.	Background.....	106
3.3.2.	Methodology and statistical analysis	108
3.3.2.1.	Influence of ROI size.....	108
3.3.2.2.	Influence of ROI Depth	109
3.3.2.3.	Influence of Gain setting	109
3.3.2.4.	Influence of focus position	111
3.3.2.5.	Influence of transducer frequency	111
3.3.3.	Results.....	112
3.3.3.1.	Sample characteristics	112
3.3.3.2.	Influence of ROI size.....	112
3.3.3.3.	Influence of the ROI depth	113
3.3.3.4.	Influence of the focus setting.....	115
3.3.3.5.	Influence of gain setting	116
3.3.3.6.	Influence of transducer frequency	118
3.3.4.	Discussion	120
3.3.5.	Testing ASQ settings	123
3.3.5.1.	Results	124
3.3.5.2.	Discussion and Conclusion.....	126
3.4.	(C) Influence of pre-defined image parameters on ASQ output	128
3.4.1.	Methodology	129

	3.4.2. Results.....	133
	3.4.3. Discussion.....	138
3.5.	(D) Applying ASQ to Images of Benign and Malignant Masses	140
	3.5.1. Methodology and Statistical analysis.....	140
	3.5.1.1. Participants.....	141
	3.5.1.2. Procedure: Recruitment and Consent.....	141
	3.5.2. Results.....	142
	3.5.3. Discussion and Conclusion.....	148
3.6.	Summary.....	152
	3.6.1. Section A.....	152
	3.6.2 Section B.....	153
	3.6.3 Section C.....	154
	3.6.4. Section D.....	155
4.	Texture Analysis Features.....	156
	4.1. Materials.....	157
	4.1.1. MaZda Software.....	157
	4.1.2. Region of interest (ROI) selection.....	158
	4.1.3. Texture features.....	160
	4.1.3.1. Grey level Co-occurrence matrix (GLCM).....	160
	4.1.3.2. Wavelet.....	161
	4.1.3.3. Tissue characterisation.....	162
	4.1.4. Data acquisition and image analysis.....	163
	4.2. Method.....	163
	4.2.1. Study Design.....	164
	4.2.1.1. Choice of Methods.....	164
	4.2.1.1.1. Sample size.....	164
	4.2.2. Participants.....	165
	4.2.3. Procedure: Recruitment and Consent.....	165
	4.2.4. Demographic Data.....	166
	4.2.5. Ultrasound.....	167
	4.2.6. Data Analysis.....	169
	4.2.6.1. Performance measures.....	171
	4.2.7. Location of Study and Access Arrangements.....	172
	4.2.8. Ethical Considerations.....	172
	4.3. Results: texture analysis.....	173
	4.3.1. Normality testing.....	178
	4.3.2. Test for significance.....	182

4.3.3. ROC and AUC	185
4.3.4. Diagnostic scoring systems	193
4.3.4.1. RMI.....	193
4.3.4.2. PMI	194
4.3.4.3. ADNEX	195
4.3.4.4. Sensitivity and Specificity	196
4.3.5. Tissue characterisation	196
4.3.6. Premenopausal group	197
4.3.6.1. PMI	203
4.3.6.2. RMI.....	203
4.3.6.3. ADNEX	204
4.3.6.4. Sensitivity and specificity	205
4.3.7. Postmenopausal group	206
4.3.7.1. PMI	211
4.3.7.2. RMI.....	211
4.3.7.3. ADNEX	212
4.3.7.4. Sensitivity and Specificity	213
4.3.8. Combination of the two texture analysis features	215
4.3.4.1. Sensitivity and Specificity	217
4.3.9. Logistic regression	218
4.4. Summary	220
(A) GLCM.....	220
(B) Wavelet.....	221
4.5. Discussion	222
5. General discussion, Conclusion, Limitations and future studies 231	
5.1 General discussion and Conclusion.....	231
5.1.1. ASQ.....	232
5.1.2. Texture analysis	234
5.2. Study limitations	238
5.3. Future work.....	239
Presentations and Posters.....	240
References	242
APPENDICES	281
APPENDIX I.....	282
APPENDIX II.....	283
APPENDIX III.....	285

APPENDIX IV	286
APPENDIX V	290

List of Tables

Table 1:	Ultrasound features used in the International Ovarian Tumour Analysis (IOTA) simple rule.....	34
Table 2:	The original Sassone scoring system for adnexal masses (Sassoneet.al.,1991).....	35
Table 3:	The DePriest scoring system (DePriest et al., 1993).....	37
Table 4:	RMI as illustrated by Prys Davies et al. (1993).....	38
Table 5:	Summary of the systemic review of preoperative methods to diagnose adnexal masses (Dodge et al., 2012).....	56
Table 6:	Pooled summary estimates of sensitivity and specificity in the new systematic review (Kaijser et al., 2014).....	57
Table 7:	Advantages of transvaginal transducer (Rumack,2011).....	91
Table 8:	Intra operator repeatability(different images).....	99
Table 9:	ASQ software repeatability (same image).....	100
Table 10:	Inter-operator mean difference and standard deviation.....	101
Table 11:	Coefficient of variation for inter-operator reproducibility.....	101
Table 12:	Summary of the influence of the ROI size on the mean of output curve using Linear Probe.....	112
Table 13:	Summary of the influence of the ROI size on the mean of output curve using Curve Probe.....	112
Table 14:	Summary of the influence of the ROI size on the mean of output curve using TV Probe.....	113
Table 15:	Summary of the influence of the ROI depth on the mean of output curve using linear probe.....	114
Table 16:	Summary of the influence of the ROI depth on the mean of output curve using curved probe.....	114
Table 17:	Summary of the influence of the ROI depth on the mean of output curve using TV probe.....	115
Table 18:	Summary of the influence of the Focus setting on the mean of output curve using linear probe.....	115
Table 19:	Summary of the influence of the focus setting on the mean of output curve using curved probe.....	116
Table 20:	Summary of the influence of the focus setting on the mean of output curve using TV probe.....	116
Table 21:	Summary of the influence of the gain setting on the mean of output curve using linear Probe.....	117
Table 22:	Summary of the influence of the gain setting on the mean of output curve using curve Probe.....	117
Table 23:	Summary of the influence of the gain setting on the mean of output curve using TV Probe.....	117
Table 24:	Influence of frequency in linear transducer.....	118
Table 25:	Influence of frequency in curved transducer.....	119
Table 26:	Influence of frequency in TV transducer.....	120
Table 27:	The difference between Toshiba model Ultrasound machines.....	125
Table 28:	Summary of the results of ASQ machine settings.....	126
Table 29:	The default settings in the ASQ software window.....	128
Table 30:	Summary of different x and y samples with their corresponding numbers of samples.....	133
Table 31:	ASQ Phantom images (homogeneous).....	134

Table 32:	ASQ tissue images (inhomogeneous).....	135
Table 33:	The significance of the difference between the phantom .images and the tissue images.....	135
Table 34:	The effect of the sample size parameter on the mean of the curve in tissue images.....	136
Table 35:	Comparison between drawing large ROI and smallest ROI with the new chosen setting on phantom images.....	137
Table 36:	Comparison between drawing large ROI and smallest ROI with the new chosen setting on tissue images.....	138
Table 37:	Shapiro-Wilk test results for normality of the data.....	144
Table 38:	Summary of mean and standard deviation of the curves (ROI including the whole mass).....	145
Table 39:	Summary of mean and standard deviation of the curves (ROI including the solid area only).....	146
Table 40:	Summary results of the significance test (ROI includes the whole mass).....	147
Table 41:	Summary results of the significance test (ROI includes only the solid area).....	147
Table 42:	Calculation of statistical performance tests. (TP is true positive, FP is false positive, FN is false negative, and FN is false negative and FN II is false negative)	171
Table 43:	Summary of histology results of 33 suspicious benign masses....	174
Table 44:	Summary of histology results of 29 malignant masses.....	175
Table 45:	Summary of the mean, median, standard deviation, and standard error of the mean of the extracted features.....	176
Table 46:	Shapiro-Wilk test for normality.....	178
Table 47:	Summary results of the significance test in all group pairs for both GLCM and wavelet.....	182
Table 48:	Tests of significant difference for the benign subgroups.....	183
Table 49:	Tests of significant difference test for malignant vs. benign subgroups.....	184
Table 50:	Area under the curve (AUC) with standard error (S.E.) associated with both GLCM and Wavelet features.....	189
Table 51:	Sensitivity and Specificity associated with GLCM and Wavelet features.....	189
Table 52:	The area under the curve (AUC) associated with both GLCM and Wavelet features in the subgroups.....	190
Table 53:	Sensitivity and specificity associated with GLCM and Wavelet features in the subgroups.....	193
Table 54:	2x2 contingency table for RMI scores.....	194
Table 55:	2x2 contingency table for PMI score.....	195
Table 56:	2x2 contingency table for the ADNEX score.....	195
Table 57:	Summary of the sensitivity, specificity, PPV, NPV and accuracy of the texture analysis features GLCM and wavelet compared to RMI, PMI and ADNEX.....	196
Table 58:	The significance test between normal ovarian tissue and malignant solid tissue.....	197

Table 59:	Summary of results of the significance test in premenopausal women.....	198
Table 60:	The area under the curve (AUC) and standard error (S.E.) associated with both GLCM and Wavelet features in premenopausal women.....	199
Table 61:	Sensitivity and Specificity associated with GLCM and Wavelet features in premenopausal women.....	200
Table 62:	2x2 contingency table for PMI score in premenopausal women.....	203
Table 63:	2x2 contingency table for RMI score in premenopausal group.....	204
Table 64:	2x2 contingency table for the ADNEX score in the premenopausal group.....	204
Table 65:	Summary of the sensitivity, specificity, PPV, NPV and accuracy of the texture analysis feature GLCM compared to RMI, PMI and ADNEX in the premenopausal group.....	205
Table 66:	Summary results of the significance test in postmenopausal women.....	207
Table 67:	The area under the curve (AUC) and standard error (S.E.) associated with both GLCM and Wavelet features in postmenopausal women.....	207
Table 68:	Sensitivity and specificity associated with GLCM and Wavelet features in postmenopausal women.....	208
Table 69:	2x2 contingency table for PMI score in postmenopausal women.....	211
Table 70:	2x2 contingency table for RMI score in postmenopausal group.....	212
Table 71:	2x2 contingency table for the ADNEX scores In the postmenopausal group.....	212
Table 72:	Summary of the sensitivity, specificity, PPV, NPV and accuracy of the Texture analysis feature wavelet compared to RMI, PMI and ADNEX in the postmenopausal group.....	213
Table 73:	Summary of the sensitivity, specificity, PPV, NPV and accuracy of the ADNEX model with only CA125 available in all three groups of population.....	215
Table 74:	2x2 contingency table for the combined two texture analysis features.....	216
Table 75:	2x2 contingency table for the combined two texture analysis features in.....	217
Table 76:	2x2 contingency table for the combined two texture analysis features in postmenopausal women.....	217
Table 77:	Summary of the sensitivity, specificity, PPV, NPV and accuracy of the texture analysis features (GLCM and wavelet) combined.....	218
Table 78:	Logistic regression model result.....	219

List of Figures

Figure 1:	Female Reproductive system.....	10
Figure 2:	(a) to 2(d) illustrate the extent of the spread for stage 1 to stage 4 of ovarian cancer.....	14
Figure 3:	examples of B-mode images of ovarian tumours.....	30
Figure 4:	Types and sub-types of texture analysis.....	62
Figure 5:	Rayleigh scattering.....	75
Figure 6:	Rayleigh pdf.....	76
Figure 7:	Normal distribution in histogram of a homogenous phantom.....	78
Figure 8:	Example of homogenous tissue histogram and non-homogenous tissue.....	79
Figure 9:	The tail in the curve produced by a discrete high amplitude signals.....	79
Figure 10:	The blue curve in the histogram.....	80
Figure 11:	The Q-Q graph in ASQ window.....	80
Figure 12:	Ultrasound machine by Toshiba.....	86
Figure 13:	How an Ultrasound scanner works.....	88
Figure 14:	Ultrasound Probe components.....	90
Figure 15:	Transvaginal Ultrasound probe.....	91
Figure 16:	ASQ window with large ROI.....	94
Figure 17:	Linear Probe.....	95
Figure 18:	Convex probe.....	95
Figure 19:	Transvaginal Probe (TV).....	96
Figure 20:	Images of RMI phantom model 403 GS LE.....	96
Figure 21:	Example of Bland-Altman plot showing the random distribution of the difference between the two operators (curve probe).....	102
Figure 22:	Example of Bland-Altman plot showing the random distribution of the difference between the two operators (linear probe).....	102
Figure 23:	Example of Bland-Altman plot showing the random distribution of the difference between the two operators (TV probe).....	103
Figure 24:	Experiment process flow illustration.....	110
Figure 25:	ASQ window with large ROI.....	113
Figure 26:	ASQ window with small ROI.....	113
Figure 27:	ASQ window with ROI at 2cm depth.....	114
Figure 28:	ASQ window with ROI at 4cm depth.....	114
Figure 29:	ASQ window with focus at 2 cm.....	115
Figure 30:	ASQ window with focus at 4 cm.....	115
Figure 31:	ASQ window with High Gain of 100 %.....	117
Figure 32:	ASQ window with low Gain of 85 %.....	117
Figure 33:	Block diagram of ASQ parameter assessment.....	131
Figure 34:	Flow chart of experiments.....	132
Figure 35:	ASQ window with the chosen set of parameters.....	133
Figure 36:	ASQ window for tissue image with default setting.....	135
Figure 37:	ASQ window for tissue image with chosen pre-set parameters.....	137
Figure 38:	Example of Q-Q plot for normally distributed data (Red curve-technique 2).....	143

Figure 39:	Example of Q-Q plot for a non-Normally distributed data (the ratio between the red and blue curve- technique 1).....	144
Figure 40:	flowchart of analysis pathways in Mazda package (adopted from (Szczypiński et al., 2009).....	159
Figure 41:	Showing computation of co-occurrence matrix with pixel of interest (Lu et al., 2006).....	161
Figure 42:	Graph showing mean value of GLCM feature for benign, malignant and simple cyst tissue.....	177
Figure 43:	Graph showing mean values of wavelet feature for benign, malignant and simple cyst tissue.....	177
Figure 44:	Q-Q plot for GLCM in the simple cyst tissue.....	179
Figure 45:	Q-Q plot for GLCM in the benign tissue.....	179
Figure 46:	Q-Q plot for GLCM in the malignant tissue.....	180
Figure 47:	Q-Q plot for wavelet in the simple cyst tissue.....	180
Figure 48:	Q-Q plot for wavelet in the benign tissue.....	181
Figure 49:	Q-Q plot for wavelet in the malignant tissue.....	181
Figure 50:	ROC curve for GLCM feature to discriminate between benign and malignant masses.....	185
Figure 51:	ROC curve for Wavelet feature to discriminate between benign and malignant masses.....	186
Figure 52:	ROC curve for GLCM feature to discriminate between cysts and benign masses.....	186
Figure 53:	ROC curve for wavelet feature to discriminate between cysts and benign masses.....	187
Figure 54:	ROC curve for GLCM feature to discriminate between cysts and malignant masses.....	187
Figure 55:	ROC curve for Wavelet feature to discriminate between cyst and malignant masses.....	188
Figure 56:	ROC curve of the GLCM feature to discriminate between malignant masses and endometriomas.....	191
Figure 57:	ROC curve of the GLCM feature to discriminate between malignant and fibroid masses.....	191
Figure 58:	ROC curve of the wavelet feature to discriminate between malignant masses and endometriomas.....	192
Figure 59:	ROC curve for the GLCM feature for the difference between malignant and benign masses in the premenopausal.....	200
Figure 60:	ROC curve for the GLCM feature for the difference between benign masses and cysts in the premenopausal group.....	201
Figure 61:	ROC curve for the wavelet feature for the difference between benign masses and cysts in the premenopausal group.....	201
Figure 62:	ROC curve for the GLCM feature for the difference between malignant masses and cysts in the premenopausal group.....	202
Figure 63:	ROC curve for the wavelet feature for the difference between malignant masses and cysts in the premenopausal	202
Figure 64:	ROC curve for the wavelet feature for the difference between malignant and benign masses in the postmenopausal group.....	208
Figure 65:	ROC curve for the GLCM feature for the difference between benign masses and cysts in the postmenopausal group.....	209
Figure 66:	ROC curve for the wavelet feature for the difference between benign masses and cysts in the postmenopausal group.....	209

Figure 67:	ROC curve for the GLCM feature for the difference between malignant masses and cysts in the postmenopausal group.....	210
Figure 68:	ROC curve of the wavelet feature for the difference between malignant masses and cysts in the postmenopausal group	210
Figure 69:	ROC for logistic regression model.....	219

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
ACS	American Cancer Society
ADNEX	Assessment of Different NEoplasias in the adneXa
AHRQ	Agency for Healthcare Research and Quality
AJCC	American Joint Committee on Cancer
AR	Auto-regressive
ASQ	Acoustic Structure Quantification
AUC	Area under the curve
BMI	Body Mass Index
CA-125	Cancer Antigen
CAD	Computer Aided Diagnosis
CoV	Coefficient of variation
CT	Computerised tomography
EUS	Endoscopic ultrasound
FIGO	International Federation of Gynaecology and Obstetrics
GLCM	Grey level co-occurrence matrix
HE4	Human Epididymis Protein 4
HRT	Hormonal Replacement Therapy
IOTA	International Ovarian Tumour Analysis group
MHz	Mega Hertz
OCS	Ovarian Crescent Sign
PDF	Probability Density Function
PMI	Pelvic Mass Index
PPV	Positive Predictive Value
PZ	Piezoelectric
RCOG	Royal College of Obstetricians and Gynaecologists
RF	Radio Frequency
RI	Resistive Index
RLM	Run length matrix
RMI	Risk Malignancy Index
ROC	Receiver operating characteristic
ROI	Region of interest
ROMA	Risk Of Malignancy Algorithm
SD	Standard deviation
TV	Transvaginal
US	Ultrasound

1. Introduction

Ovarian cancer is the second most common gynaecological malignancy; however, it remains the leading cause of death among these diseases (Givens et al., 2009). In spite of the diagnostic and therapeutic advances in the care of women with ovarian cancer, the overall five-year survival rate remains unchanged. The reason for this is that most cases are diagnosed in the late stages of the disease, when the five-year survival rates falls below 20% (Dutta et al., 2010a).

CA125 is a commonly used tumour marker to assess ovarian cancer. However, the usefulness of this marker is limited due to lack of sensitivity, as it is elevated in only 50% of patients with stage 1 cancer (Ronco et al., 2011) and has poor specificity because it is elevated in numerous benign conditions such as endometriosis, adenomyosis and pelvic inflammatory disease, as well as several non-gynaecological conditions such as diverticulitis, liver and heart failure, and in cancer of the pancreas, breast, bladder and liver (Ronco et al., 2011).

Researchers have combined ultrasound with CA125 to increase the sensitivity and specificity of differentiating between benign and malignant masses. As a result, the Risk of Malignancy Index (RMI) was created (Jacobs et al., 1990). It became recommended by the Royal College of Obstetricians and Gynaecologists. However, it has a sensitivity of 89% and a specificity of 73% (RCOG, 2003).

A study in 2005 concluded that ultrasound has a high false-positive rate in the differential diagnosis of adnexal malignancies, even with using several scoring systems (Guerriero et al., 2005).

Other researchers have used Doppler to differentiate between benign and malignant ovarian masses to improve the specificity of ultrasound. Unfortunately, according to most of these studies, this approach does not add significant useful information, with a reported accuracy of only 35%- 88% (Kinkel et al., 2000, Gentry-Maharaj and Menon, 2012).

Others have tested the usefulness of using four markers as a combined test for early-stage ovarian cancer detection and showed a sensitivity and specificity of 91.3% and 88.5% respectively (Gentry-Maharaj and Menon, 2012). These panels, nevertheless, have yet to be validated widely in clinical trials.

Other preoperative methods to differentiate benign from malignant ovarian masses have been developed and tested recently, such as ultrasound contrast agents (Fleischer et al., 2012, Dutta et al., 2010a), 3D imaging (Huchon et al., 2012) and using CT (Palma et al., 2012) or MRI imaging. However, they all have limitations, including limited availability, high cost, difficulty with the identification of small tumour deposits, cancer extension and the distinction between benign and pathological lesions (Iyer and Lee, 2010, Dutta et al., 2010a).

In recent years, objective diagnostic methods have been proposed to overcome the limitations of subjectivity and operator dependence. However, there is no reliable

technique available at present. Hence, a new objective method is desired to address the above-mentioned issues which will contribute in patient management.

Texture analysis is a technique for evaluating the structure within an image. In digital imaging, texture analysis is the analysis of the distribution of grey level values across the pixels of a given region of interest. The variation in intensity reflects some physical variation in the underlying structure, as explained by (Szczyński et al., 2009); moreover, the image texture of medical images describes the internal structure of human tissue or organs. It can also describe pathological changes (Xian, 2010). The texture analysis of ultrasound images relies on the principle that, when disease occurs and starts to affect the structure of the tissue, the tissue should reflect a different ultrasound signal, which will in turn cause the texture feature to give a different value to normal tissue (Morris, 1988)

Texture analysis techniques have been successfully applied to various types of tissues and organs (Michail et al., 2007, Giger et al., 2008), including the carotid artery (Coleman et al., 2005), breast (Alacam et al., 2003, Chen et al., 2002, Huang et al., 2008, Ramaraj and Raghavan, 2011), heart (Vince et al., 2000), thyroid gland (Smutek et al., 2003), prostate (Basset et al., 1993), pancreas (Das et al., 2008) and liver (Lee et al., 2003, Xian, 2010, Vicas et al., 2011, Kumar et al., 2012). However, it has not yet been widely tested on ovarian tissue.

A preliminary study has demonstrated the usefulness of texture analysis in differentiating ovarian lesions (Hamid et al., 2011). Therefore, objective differentiation between benign and malignant ovarian tissue through texture analysis

on a larger sample would be beneficial by validating the accuracy of the method and therefore decreasing the rate of unnecessary surgery.

This is a prospective cross-sectional study of patients with adnexal masses to validate texture analysis as a method of objectively differentiating ovarian lesions to overcome the subjectivity and operator-dependent limitations of Ultrasound.

1.1. Research hypothesis

Based on recent developments in ultrasound imaging, computer technology and extensive research on texture analysis with a focus on medical images, it is hypothesized that the texture analysis technique can be used to characterize and quantify ovarian tissue based on B-mode image texture.

The motivation for this technique is clear: since the output of texture analysis of an image can be expressed numerically, it provides a quantitative means of image description, which could help in reducing subjectivity. This technique would also overcome the reproducibility issue encountered as a result of the subjective interpretation of the image. Combining human skills with results from computers to aid diagnosis is expected to improve the overall diagnosis of ovarian masses.

1.2. Aims and objectives

Study aims: To assess prospectively the diagnostic performance of texture analysis on grey-scale transvaginal ultrasound images in discriminating between benign and malignant adnexal tumours and to compare it to other widely used scoring systems.

The primary end-point is the accuracy of diagnosing ovarian cancer when compared with histology.

The secondary end-point is an attempt to correctly classify the pathology and compare the sensitivities and specificities of GLCM and Wavelet tissue characterization techniques compared with other widely used diagnostic models.

1.2.1. Objectives:

In achieving the above aim, the following objectives have been set:

- To estimate the accuracy of texture analysis in differentiating ovarian lesions against histology results.
- To compare texture analysis performance with other widely used scoring systems.
- To improve the sensitivity and specificity of transvaginal ultrasound in the diagnosis of ovarian lesions through objective assessment of the ultrasound images.

1.3. Thesis structure

This study is divided into two parts. The first part focuses on Acoustic Structural Quantification (ASQ). This part can be further divided into four main sections: section A will include a definition of the new term followed by background information and the phantom study in which repeatability and reproducibility will be tested; then, section B will study Influence factors such as ROI size, ROI depth, Focus, Gain setting, and frequency on ASQ. Then, in section C, the influence of pre-defined image parameters on ASQ output will be discussed. Lastly, section D will demonstrate the application of ASQ on images of benign and malignant masses.

The second part focuses on texture features and consists of materials and methods used in the study, results and figures, a discussion of the main results and then the conclusion of the study.

This thesis is structured into six main chapters:

Chapter 1: this chapter gives an overview of the whole thesis.

Chapter 2: this chapter critically evaluates current ultrasonic techniques used to diagnose ovarian cancer and related issues. Alternative techniques are also reviewed in this chapter.

Chapter 3: this chapter introduces the utilization of ASQ technique for identification of ovarian cancer tissues.

Chapter 4: this chapter focuses firstly features on explaining two of texture analysis methods, namely GLCM and wavelet, then applies these methods to ultrasound images of ovarian masses.

Chapter 5: this chapter contains the general discussion of the study results of both ASQ and texture analysis in ovarian cancer early diagnosis.

Chapter 6: this chapter provides a summary and conclusion of the study and mentions its limitations, as well as recommending possible future work.

2. Literature Review

2.1. Background

Ovarian cancer is the second most common gynaecological malignancy in the developed countries and is the fifth leading cause of death due to cancer among women (McDonald et al., 2010, Enakpene et al., 2009, Leitzmann et al., 2009, Jemal et al., 2009, Ulusoy et al., 2007, Goff et al., 2007, Varras, 2004, Morgante et al., 1999). Globally, the average lifetime risk of ovarian cancer is about 1 in 70 women, and cancer is seen very rarely before the age of 40 (Loubeyre et al., 2012).

About 21,290 new cases of ovarian cancer are expected in 2015. Moreover, an estimated 14,180 deaths are expected in the USA in 2015. It causes more deaths than any other cancer of the female reproductive system. The majority of cases (63%) are diagnosed in the late stages, with a five-year survival rate of 27%, in comparison with a 72% survival rate when detected in the early stages (American Cancer Society, 2015). In the European Union, an estimated 44,149 new cases were diagnosed and 29,770 deaths from ovarian cancer occurred in 2012 (International Agency for Research on Cancer, 2012).

In the United Kingdom, just over 7,000 new ovarian cancer cases are diagnosed every year (Cancer Research, 2012). According to Cancer Research UK, of all women diagnosed with ovarian cancer, 40% will live at least five years after diagnosis: the 30% diagnosed with stage I cancer will have a 90% five-year survival rate, while the 15% diagnosed with late stage cancer will have only a 6% 5-year survival rate (Cancer Research, 2012). Ovarian cancer becomes more common with increasing age.

Three-quarters of new ovarian cancer cases in the UK are diagnosed in women aged 55 and over.

As can be seen from the above, early diagnosis is a vital factor for prognosis. Ultrasound is considered the main imaging modality for ovarian cancer triage (Kinkel et al., 2000). However, there is concern about the reproducibility of the diagnosis, mainly due to the subjective nature of the interpretation of the images and its dependency on the experience of the observer. Other methods of diagnosis are tumour marker analysis as well as other imaging modalities such as magnetic resonance imaging (MRI) and computed tomography (CT).

2.1.1. Ovarian cancer

Normally, normal cells grow and divide to form new cells. When uncontrolled and abnormal growth and division of the cells occurs, it causes cancer. A mass of tissue of these extra cells will form and is called tumour, which can be benign (non-cancerous) or malignant (cancerous). Benign tumours do not invade the tissues around them and do not spread to other parts of the body (metastasis). Malignant tumours, however, have the ability to invade and destroy other tissues surrounding them. Cancer cells are able to metastasise via the blood stream or the lymph system and spread to other parts of the body. Ovarian cancer refers to cancer of the ovaries (the female reproductive organs).

2.1.1.1. Ovaries: location and appearance

The ovaries are a pair of oval-shaped organs that normally measure 2-4 cm in diameter and are located in the pelvis, on each side of the uterus. Two individual

functions of the ovaries are: (i) to produce eggs, and (ii) to produce the female hormones (oestrogen and progesterone). (See figure 1)

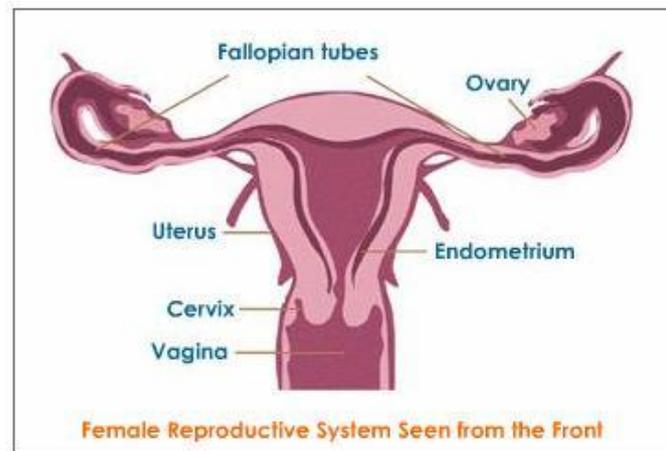


Figure 1: Female Reproductive system

Generally, uterine location influences the position of the ovaries. Normal ovaries are usually identified laterally or posterolaterally to the uterus. In cases of retroverted uterus, the ovaries tend to be located laterally and superiorly, near the fundus. Because of their variable position, superiorly or extremely laterally placed ovaries may not be visualised using the Transvaginal approach because they are out of the field of view.

The ovaries are ellipsoid in shape, with their craniocaudal axes paralleling the internal iliac vessels, which lie posteriorly and serve as a helpful reference.

On sonography, the normal ovary has a relatively homogenous echo texture with a central more echogenic medulla; well-defined, small anechoic or cystic follicles may

be seen peripherally to the cortex. The appearance of the ovaries changes with age and with the phase of the menstrual cycle (Rumack, 2005, 2011).

Because of variability in shape, ovarian volume has been considered the best method for determining ovarian size. Volume measurement is based on the formula for the prolate ellipse (see equation 1).

(width × height × thickness × 0.523) or (d1 × d2 × d3 × 0.532)

Equation 1: Ovarian volume

Adult measurement of the ovary can be as large as 22 cm³. Following menopause, the ovary decreases in size with increased age. Due to its smaller size and lack of follicles, the postmenopausal ovary may be difficult to visualise sonographically. Mean ovarian volume ranges from 1.2 cm³ to 5.8 cm³. Ovarian volume > 8 cm³. is definitely considered abnormal (Rumack, 2005). Ovarian masses can be classified as simple cysts, complex cysts, or solid masses. Simple cysts are typically unilocular of any size, echo free and thin-walled. Thin septa (<3mm) may be present. The probability that a simple cyst will be malignant in a peri- or postmenopausal woman is exceptionally low (<1%) (Ekerhovd et al., 2001, Sharma et al., Published Online 2011). Complex cysts have echogenic content, papillary formation, thick inner septa (>3mm), thick walls and might contain solid components. Even though most complex cysts are benign when examined histologically, there are no imaging features that can differentiate benign from malignant tissue in these lesions. The presence of a solid intracystic area increases the likelihood of malignancy. Solid masses usually raise the suspicion of malignancy, particularly when extra-pelvic ascites are present (Woodward et al., 2004).

Generally, ovarian cancer can develop at any age, yet it is most likely to occur in older age, between 40 and 65 years (Crum, 2004, Loubeyre et al., 2012). Serous adenoma, mucinous adenoma and Brenner tumours are all examples of benign ovarian tumours, whereas serous adenocarcinomas and mucinous adenocarcinomas are examples of malignant ovarian tumours which originate from epithelial cells. According to (Jeong et al., 2000), epithelial cell origin tumours account for 85% of all ovarian cancer cases. Commonly, epithelial tumours occur in peri-menopausal and post-menopausal women, with a mean age of 55 years (Russell, 1994).

2.1.1.2. Ovarian cancer: types

Ovarian cancer is grouped into three major types according to the tissue of the ovary it originated from: (i) epithelial tissue, (ii) germ cell and (iii) sex cord-stromal cell. The most common type is epithelial, accounting for 85% of cases. Epithelial tumours are rare before puberty. Their frequency increases with age and reaches a maximum at 60 years of age (Jeong et al., 2000). Germ cell tumours are less common, accounting for 5-10% of cases. They usually affect younger women, where the peak incidence is in the early 20s. Sex cord-stromal tumours are rare, accounting for only less than 5% of the total cases. They can be found in all age groups. Sex cord-stromal tumours originate in the connective cells that hold the ovaries together and produce female hormones.

2.1.1.3. Ovarian Cancer: Staging

Staging describes the extent or spread of cancer to other parts of the body at the time of diagnosis. Proper staging is essential in determining the type of therapy, such as surgery and chemotherapy planning, and in predicting the survival rate. Several

different staging systems are used to classify tumours. The TNM staging system developed by the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC) evaluates tumours in three ways: the extent of the primary tumour (T), the presence or absence of regional lymph node involvement (N) and the presence or absence of distant metastases (M). After the T, N and M categories are assigned, they can be further grouped into stages I, II, III or IV, with stage I being early and stage IV being advanced disease according to the FIGO staging system developed by the International Federation of Gynecology and Obstetrics (Fischerova, 2011).

The document that explains the FIGO ovarian cancer staging is available online at: (http://www.figo.org/files/figo-corp/docs/staging_booklet.pdf). Briefly, the four stages of ovarian cancer are as follows and the illustrations are given in figure 2(a) to 2(d):

Stage 1: the cancer is limited to one or both ovaries and has not spread to other organs.

Stage 2: the cancer can be found outside of the ovary, but has not spread further than the pelvic region (uterus, bladder, lower intestine).

Stage 3: the cancer is limited to the peritoneal cavity.

Stage 4: the cancer has spread beyond the abdomen to areas such as the liver, lungs and brain.

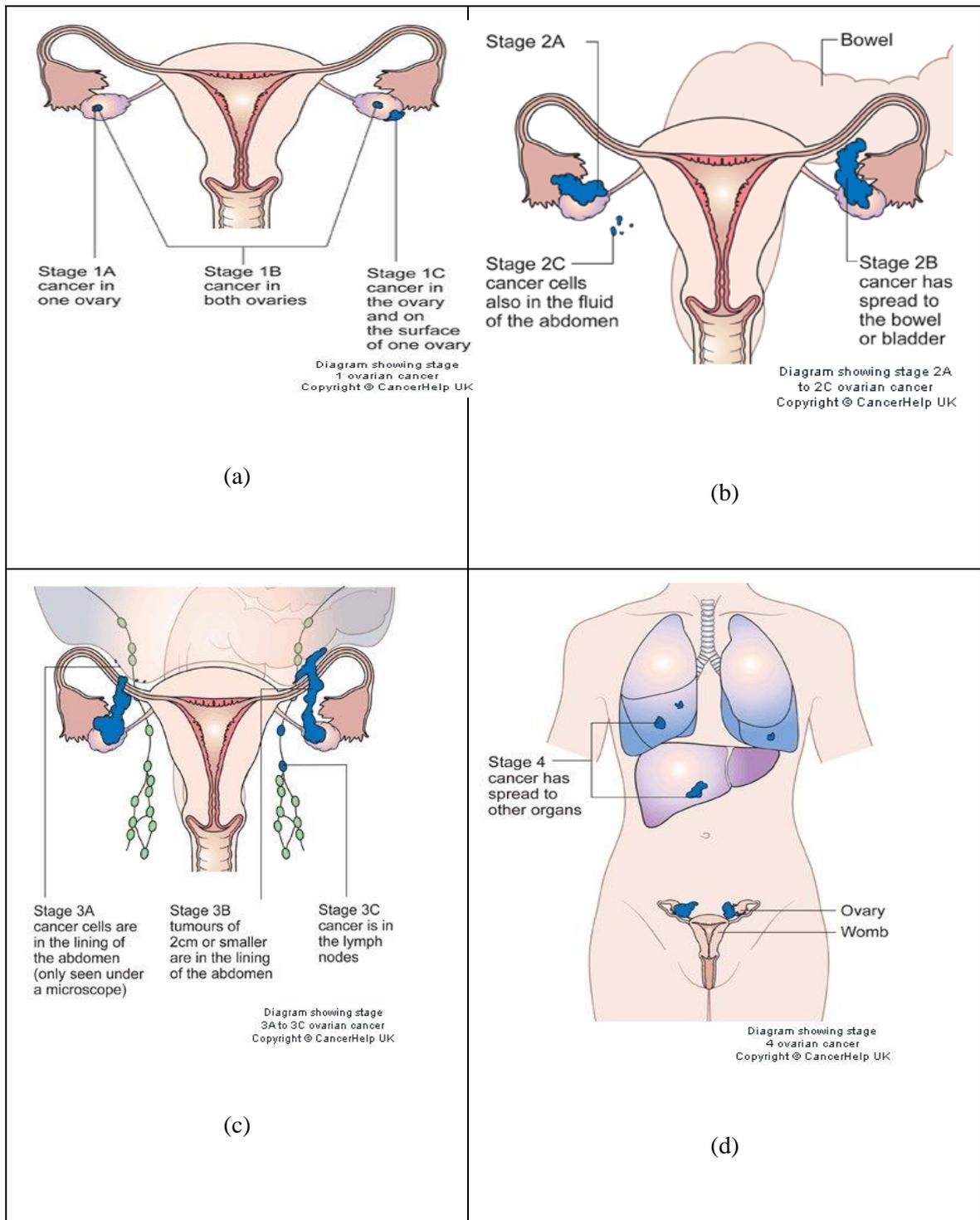


Figure 2:(a) to 2(d) illustrate the extent of the spread for stage 1 to stage 4 of ovarian cancer (adopted from: <http://www.cancerresearchuk.org/cancer-help/type/ovarian-cancer/treatment/stages-of-cancer>)

2.1.1.4. Risk Factors

Although extensive research has been carried out, the aetiology of ovarian cancer is poorly understood. Nevertheless, the two most prominent factors associated with the risk of developing ovarian cancer are the increase in age and the presence of certain gene mutations.

Several factors, such as oral contraceptive use, parity and breastfeeding are believed to be associated with decreased risk (Permuth-Wey and Sellers, 2009). Body mass index (BMI), height, family history of ovarian cancer and prolonged use of postmenopausal hormonal therapy (HRT) are associated with increased risk of ovarian cancer (Givens et al., 2009, Permuth-Wey and Sellers, 2009, Schouten et al., 2008, McBee et al., 2007, Olsen et al., 2007, Engeland et al., 2003).

2.1.1.4.1. Age

Incidence of ovarian cancer increases with age, with a median age at diagnosis of 63 years. Over 80% of cases occur after the age of 45 years. Okugawa and colleagues studied the relationship between age, histological type and the size of ovarian tumours and found that tumours in older patients are more likely to be malignant (Okugawa et al., 2001). Moreover, McBee and colleagues considered patient age as the most important predictor of whether a mass is likely to be malignant or benign (McBee et al., 2007). In addition, more than one-third of cases of ovarian cancer occur in women older than 65 years of age (McBee et al., 2007, Varras, 2004). In women under the age of 45 years, the risk that an ovarian mass will be a primary malignancy is only one in fifteen (Varras, 2004). It has been reported that the risk of ovarian cancer

increases progressively with age, with a higher risk after the menopause (Givens et al., 2009).

2.1.1.4.2. Family History

One of the most significant risk factors for ovarian cancer is a family history of breast, colon or ovarian cancer, especially if multiple first degree relatives (mother or sister) are diagnosed with ovarian cancer. Similarly, women with a previous personal history of breast cancer have double the risk of ovarian cancer, and the risk is increased nearly four-fold for women diagnosed with breast cancer before the age of 40 (Permuth-Wey and Sellers, 2009, Givens et al., 2009).

2.1.1.4.3. Height

When taking the height of women into account as a risk factor, Schouten et al, (2008) found in their review of twelve prospective cohort studies from North America and Europe that height was not associated with an increased risk of cancer in postmenopausal women. Conversely, a Norwegian cohort of 1.1 million women with 7,882 cases of ovarian cancer reported a positive association between taller women (more than 1.75 m) and ovarian cancer compared with women between 1.60 and 1.64 m in height (Engeland et al., 2003), and a recent Dutch study had similar results, finding a positive association in postmenopausal women between height and risk of ovarian cancer (Schouten et al., 2008).

Gunnell conducted a study explaining that height on its own does not cause cancer but possibly acts as an indicator for some other exposure factors such as genetics, energy intake in early life and exposure to sex and growth hormones (Gunnell, 2001). This

concluded with Engeland and colleagues (2003), who demonstrated that height may be a sign that early life situations are interrelated with the risk of cancer.

2.1.1.4.4. Body Mass Index (BMI)

Studies that observe a relationship between current BMI and ovarian cancer risk in postmenopausal women are questionable. A meta-analysis in 2007 reported a 16% increase in risk for overweight women with BMI 25-29.9 kg/m² and a 30% increased risk for obese women with BMI \geq 30 kg/m² in comparison to normal weight women with BMI 18.5-24.9 kg/m² (Olsen et al., 2007).

The majority of studies have found no relationship between current BMI and increased risk of ovarian cancer in postmenopausal women. However, they showed a positive correlation of early adulthood high BMI (age 18-20 years) and increased risk of ovarian cancer. Their findings suggested that long term, consistent excess weight represents a major risk factor for ovarian cancer, especially with no family history of ovarian cancer, whereas high BMI with positive family history of ovarian cancer will not increase the risk (Leitzmann et al., 2009, Olsen et al., 2008, Lubin et al., 2003, Engeland et al., 2003, Rodriguez et al., 2002). This was supported by others, who concluded that being overweight or obese in premenopausal years (<50 years) of age is associated with an increased risk of ovarian cancer when women become postmenopausal (Permuth-Wey and Sellers, 2009). One noteworthy limitation from the Leitzmann et al. (2009) study was the limited ethnic/racial diversity of participants (Leitzmann et al., 2009). Other studies disagreed with this and concluded that there is no association between BMI in early adolescence and the risk of cancer (Permuth-Wey and Sellers, 2009, Schouten et al., 2008).

2.1.1.4.5. Hysterectomy

Numerous studies have investigated whether hysterectomy or other procedures have an effect on the risk of ovarian cancer (Rosenblatt and Thomas, 1996, Green et al., 1997, Kreiger et al., 1997, Miracle-McMahill et al., 1997, Risch et al., 1994, Hankinson et al., 1993). All of them have identified a reduced risk of ovarian cancer associated with either hysterectomy or tubal ligation (without oophorectomy). Furthermore, tubal ligation has a protective effect on ovarian cancer with an estimated reduced risk of between 18% and 70%. Although it is undetermined how these procedures reduce the risk of cancer, it has been proposed that the risk of ovarian cancer may decrease due to the decrease in blood flow to the ovaries (Permuth-Wey and Sellers, 2009).

2.1.1.4.6. Use of Hormone Replacement Therapy (HRT)

Hormones such as oestrogens and progesterone are believed to be involved in promoting ovarian carcinogenesis (Permuth-Wey and Sellers, 2009).

A recent study showed an increased risk of ovarian cancer among HRT users. It was reported that the risk increased by 30% regardless of duration of use, route of administration or type of hormone used, even with a short duration of use (0-4 years) (Mørch et al., 2009). In 2009, Permuth-Wey and Sellers illustrated that both current and past users of HRT of five or more years had a significantly higher risk compared with women who had never used any type of HRT. They concluded that the considerable increase in the risk of ovarian cancer was associated with duration rather than with status of use. They supported their findings with the UK Million Women study results, which found that the incidence of ovarian cancer increased in current

HRT users with (prolonged) duration of use but did not differ significantly by type of hormone used and that past users were not at increased risk, since the risk diminished two years after discontinuing the treatment (Permuth-Wey and Sellers, 2009). Others (Greiser et al., 2007) found similar results when available studies were reviewed and further determined that the risk is greater in European studies than in North American studies.

Other factors that are alleged to decrease the risk are:

Parity: the risk of ovarian cancer is lower in women who have children compared to women who have never had children (Modan et al., 2001).

Breastfeeding also reduces the risk of ovarian cancer in women who breastfed, compared to those who have never breastfed (Luan et al., 2013).

Infertility: null gravid women who have been attempting to get pregnant for more than five years have an increased risk compared to women who have been trying to conceive for less than a year (Rossing et al., 2013).

Oral contraceptives usage reduces the risk of ovarian cancer. (Modan et al., 2001)

Web reference: <http://info.cancerresearchuk.org/cancerstats/types/ovary/riskfactors/>

2.1.1.5. Symptoms

According to (Chan and Selman, 2006), the symptoms of ovarian cancer are non-specific. However, Goff and colleagues studied the possibility of early detection of ovarian cancer through the development of an ovarian cancer symptom index. They suggested, from their results, that specific symptoms such as abdominal/pelvic pain, increased abdominal size/bloating, or difficulty eating/feeling full quickly that are

experienced >12 times per month over a one-year duration should raise the suspicion of ovarian cancer (Goff et al., 2007). This is consistent with a study by Smith et al. (2005), who added urinary urgency or frequency to the symptom profile. Other studies agreed, such as (Vine et al., 2003) and (Olson et al., 2001), as quoted by (Cancer Research, 2012). These studies are in agreement with another (Hamilton et al., 2009) which suggested that ovarian cancer should not be called the silent killer any longer because of the symptoms mentioned in previous studies. Moreover, their study found that there were only minor differences in documented symptoms between cases with early and later stages of the disease.

In another review (McBee et al., 2007) which studied ovarian masses that require intervention, it was pointed out that the best predictors of malignant tumours are a combination of factors that include patient age, family history, menopausal status, symptoms, findings on clinical examination, ultrasound imaging findings and serum CA125 level. They listed the symptoms that suggest malignancy in women with pelvic mass as abdominal pain/bloating, abnormal vaginal bleeding or discharge, change in consistency of stool, decreased appetite, frequent urination, increased abdominal girth, nausea or vomiting and significant weight loss. They added that other clues could be found on physical examination, such as ascites or upper abdominal mass, a mass that is large, firm, irregular or fixed and nodularity on rectal examination.

Evidence shows that to differentiate between women who are experiencing symptoms related to ovarian cancer and women who are not, the frequency, persistency, severity

and new onset of these symptoms might help (Bankhead et al., 2008, Goff et al., 2007).

Other than that, loss of appetite, nausea, lower back pain and shortness of breath may also be an indication of ovarian cancer. Nevertheless, it is important to emphasize that these symptoms could be due to other diseases.

In a more recent study in 2012, the author discussed the possibility and the feasibility of using computer assisted diagnosis (CAD) to identify the symptoms of ovarian cancer by creating a predictive algorithm (Hamilton, 2012).

2.1.2. Methods of Preoperative Evaluation

Ovarian cancer is diagnosed at an advanced stage in most patients due to the non-specific nature of the symptoms and signs of this disease. Several methods have been used for the preoperative evaluation of adnexal masses, such as gynaecologic examination, blood tests for serum levels of CA-125 (Van Calster et al., 2007), imaging procedures such as subjective evaluation of grey-scale and Doppler ultrasound findings by an experienced examiner (pattern recognition), CT/ MRI, laparoscopy and laparotomy, as well as patient characteristics such as menopausal status or age.

The only way to confirm the diagnosis is to remove a tissue sample from the mass and examine it under a microscope: this is called histology (Marshall, 2008, American Cancer Society, 2015).

Before ultrasound became extensively available, the discovery of a palpable ovary or pelvic mass in a postmenopausal woman was considered an indication for surgery (Valentin, 2000). Therefore, it is of great importance that the imaging modality has a high sensitivity as well as reliable characterization of the lesions to allow accurate and consistent diagnosis. This is particularly vital since the symptoms of ovarian cancer are non-specific, and the patient management and treatment depend on tumour staging.

2.1.2.1. Tumour markers

Tumour markers are substances either produced or released by tumour cells or host cells, which point out that a tumour is present if identified in serum or other biological fluids. Preferably, a marker should be high in both sensitivity and specificity. They have the advantage of being non-invasive, quick, widely available and relatively cheap (Agarwal et al., 2011, Hellstrom and Hellstrom, 2008).

2.1.2.1.1. Serum CA125

CA stands for cancer antigen. CA 125 is a protein that is a so-called tumor marker or biomarker, which is a substance that is found in greater concentration in tumor cells than in other cells of the body. In particular, CA 125 is present in greater concentration in ovarian cancer cells than in other cells. It was first identified in the early 1980s. CA 125 is often measured as a blood test (Dong et al., 2008).

Although CA125 is the most well-known and commonly used serum tumour marker in Gynaecology (Moore and MacLaughlan, 2010, McBee et al., 2007, Gadducci et al.,

2004), the role of this biomarker in particular in the diagnosis of ovarian cancer is controversial.

The usefulness of this marker is limited due to lack of sensitivity, as it is elevated in only 50% of patients with stage 1 cancer (Ronco et al., 2011) and in 75-90% of patients with advanced disease (Moss et al., 2005), as well as having poor specificity because it is elevated in numerous benign conditions such as endometriosis, adenomyosis and pelvic inflammatory disease as well as several non-gynaecological conditions such as diverticulitis, liver and heart failure, pancreatic cancer, and breast, bladder and liver cancer (Duffy et al., 2005, Ronco et al., 2011).

Two studies based on the International Ovarian Tumour Analysis (IOTA) data by (Timmerman et al., 2007) and (Van Calster et al., 2007) examined the importance of CA125 in comparison with ultrasound, clinical information, and pattern recognition respectively. The first study suggested that when combining ultrasound with certain clinical information, the addition of serum CA125 level does not improve the preoperative classification of a mass or redirect management in either premenopausal or postmenopausal women.

The other study found that pattern recognition by an experienced examiner was superior to measuring the level of serum CA125 in differentiating between benign and malignant adnexal masses. However, the later study revealed a bias because serum CA125 was more likely to be measured in women with suspected malignant masses only; moreover, 24% of benign cases did not have a serum CA125 done and were diagnosed confidently through pattern recognition alone. These results were agreed in

a later study that found that Transvaginal ultrasound done by an expert examiner is superior to CA125 analysis in diagnosing ovarian cancer (Alcázar and Guerrero, 2011).

Hartman and colleagues carried out a study to assess ultrasound criteria and CA125 as predictive variables of ovarian cancer. They described their work as the first study to attempt to reproduce the simple rule established by (Timmerman et al., 2008) and apply them outside the European centres of the IOTA. They concluded that the simple rules allow the correct classification of ovarian masses and added that CA125 measurement can slightly improve the specificity of ultrasound examination when the scan is suggestive of malignancy; however, CA125 measurement should not be used alone in diagnosing ovarian tumours (Hartman et al., 2012).

Moreover, a study commenced by (Givens et al., 2009) suggested that CA125 levels should be checked for postmenopausal women with adnexal masses to guide treatment options, but should not be used as a screening tool or when a mass is not identified. This result was supported later by the recent IOTA recommendation in 2013 for clinical practice, where measurements of serum CA125 were not necessary for the characterization of ovarian pathology in premenopausal women. It also showed that measurements of serum CA125 are unlikely to improve the performance of experienced ultrasound examiners decision, even in the postmenopausal group (Kaijser et al., 2013).

2.1.2.1.2. Other Biomarkers

In a recent study, six different serum markers were tested for their elevation in patients with ovarian cancer. This study suggested that human epididymis protein 4 (HE4) is the best marker for use as a secondary screening test. The HE4 had a higher sensitivity in all stages of type 2 ovarian cancer and a lower sensitivity in early stages of type 1, including stage 1 and 2, when compared with Transvaginal ultrasound. However, these results need to be further validated and confirmed by an independent group that the measurement of HE4 outperforms Transvaginal ultrasound as a screening test (Urban et al., 2011). Similarly, a study done by Agarwal and colleagues, focusing on potential markers for detecting ovarian cancer, found that HE4 had a high sensitivity and specificity of 90% and 77.6% respectively. Moreover, it possessed the highest sensitivity in detecting stage 1 ovarian cancer, as well as fewer false positive results, especially in non-malignant ovarian diseases (Agarwal et al., 2011).

Studies focusing on the potential use of HE4 as a biomarker of ovarian cancer suggest that it is elevated in over 50% of ovarian cancer patients with low CA125. In addition, HE4 has a greater sensitivity than CA125 in early stage ovarian cancer, as well as greater specificity in comparison with benign ovarian masses (Moore and MacLaughlan, 2010). Based on these findings, a scoring model was developed by Steven Skates and colleagues, called the Risk of Ovarian Malignancy Algorithm (ROMA), which uses measurements of CA125 and HE4, combined with menopausal status, in order to assign high or low risk of malignancy in women with pelvic mass (Moore and MacLaughlan, 2010).

Following this, in 2012, a study was commenced to compare CA125, HE4, ROMA and RMI in the classification of ovarian masses. The four methods demonstrated similar levels of accuracy in their ability to differentiate adnexal masses, with RMI having the lowest sensitivity and HE4 the best overall sensitivity for evaluation of malignant tumours (Anton et al., 2012). (RMI will be explained in section 2.1.2.2.1.3. Risk of Malignancy Index, later in this chapter)

More recently, ROMA was evaluated in a prospective, multicentre trial involving 472 patients and provided a sensitivity of 93.8% and specificity of 74.9%. It performed well in premenopausal women, with a sensitivity of 100% and specificity of 74.2%. In addition, the study demonstrated a clear benefit of ROMA to ovarian cancer patients in terms of mortality and morbidity (Nolen and Lokshin, 2013).

Another study focused on soluble mesothelin-related proteins (SMRP) and HE4. The results showed a comparable sensitivity of HE4 with CA125 but with higher specificity. Furthermore, the authors concluded that measuring SMRP and HE4 in serum could be useful for patients with ovarian cancer, as they complement CA125 for diagnosing and monitoring patients. Nevertheless, the need for prospective studies to establish the clinical relevance of these findings is obvious (Hellstrom and Hellstrom, 2008).

Another biomarker, called human Kallikrein 10 (hK10), was evaluated for ovarian cancer diagnosis and prognosis by (Luo et al., 2003). In this study, it was demonstrated that hK10 could identify a significant number of patients who were missed when using CA125 analysis alone. Moreover, when combining hK10 with

CA125, the achieved sensitivity was 73%, which is superior to hK10 (55%) or CA125 (60%) alone with the same specificity of 90%. Thus, CA125 and hK10 can be combined to increase the diagnostic sensitivity of each of the biomarkers alone (Luo et al., 2003). Similarly, another study found that hK10 can be utilized in many malignancies but lacks sufficient specificity or sensitivity to be clinically useful when used alone (Dutta et al., 2010b).

Despite recent advances in ovarian cancer biomarkers research, no simple blood test with the required sensitivity and specificity has yet been clinically validated (Dutta et al., 2010b). This could be explained by the complexity and heterogeneity of ovarian cancer. In other words, it is doubtful that a single biomarker will be able to detect all subtypes and stages of the disease with high specificity and sensitivity (Gagnon and Ye, 2008).

2.1.2.2. Ultrasound

Ultrasonography is, at present, the most widely used diagnostic imaging technique for the differential diagnosis of adnexal masses. It has been used in medical imaging for over half a century (Hangiandreou, 2003), and it is commonly considered as the preferred imaging modality in the study of the female pelvis (Derchi et al., 2001). It is a non-invasive procedure, with relatively low cost, and is well accepted by most women. Visualization of normal ovarian function and ovarian masses has improved since the introduction of the transvaginal transducer (Twickler and Moschos, 2010).

Ultrasound is currently one of the most significant, extensively used, and valuable imaging modalities in medicine (Hangiandreou, 2003). According to Kinkel et al.

(2000), it is the imaging modality of choice in the evaluation of suspected adnexal masses; in addition, ultrasound is the main triage method for ovarian cancer prior to treatment. This was supported by a later study by Togashi (Togashi, 2003).

The main downside of ultrasound is that the accuracy of ultrasound scanning is greatly dependent on three variables: operator, equipment and patient. To become an ultrasound expert in gynaecology requires extensive practical proficiency, which is not acquired easily by every person. To examine patients effectively, a high-end ultrasound machine equipped with sensitive Doppler and endocavitary and trans-abdominal probes are needed. Another significant variable in the accuracy of ultrasound is the patient. Although there has been considerable improvement in scanning technology, there are nevertheless limitations in some cases such as obese patients and postoperative adherent intestinal loops causing acoustic shadowing (Fischerova, 2011).

B-Mode or brightness mode is a two-dimensional ultrasound image display composed of bright dots representing the ultrasound echoes. The brightness of each dot is determined by the amplitude of the returned echo signal (Rumack, 2005).

Doppler ultrasound is used to detect the motion of blood. It is based on the concept of Doppler shift, which is the change in frequency for a reflector moving relative to the source. When reflector is moving toward the transducer waves are closer together gives high frequency higher than the transmitted one. The difference between the transmitted and the received frequency is called Doppler shift (Horskin, 2010).

Ultrasound-based assessment for ovarian mass characterizations can be divided into the following techniques: morphologic information, which is based on B-mode images, and blood flow information, which is based on Doppler imaging.

The level of performance of ultrasound examination in ovarian cancer diagnosis has been studied and reported to have a wide range. Kinkel et al. (2000), for example, reported that the accuracy of ultrasound is 65%- 94% for B-mode ultrasound and 35%- 88% for colour Doppler flow imaging. On the other hand, the Royal College of Obstetricians and Gynaecologists (RCOG), in its document “Guideline No. 34: ovarian cysts in postmenopausal women”, referred to sensitivity of 89% and specificity of 73% of ultrasound performance when using a morphology index in (DePriest et al., 1994) study (RCOG, 2003)

(Kinkel et al., 2000) assert that although extensive studies have been done on ultrasound techniques and they have been found to be the best means of lesion characterization, this approach remains inconclusive. Nevertheless, it has been demonstrated that ultrasound techniques that combine grey-scale ultrasound morphology assessments with tumour vascularity imaging information (colour Doppler flow imaging) in a diagnostic system are significantly better in ovarian lesion characterization compared to using power Doppler, colour Doppler flow imaging, or grey-scale ultrasound morphologic imaging alone (Kinkel et al., 2000, Togashi, 2003).

2.1.2.2.1. B-mode (brightness-mode) Ultrasound: (see figure 3)

B-mode ultrasound uses morphological features to diagnose ovarian cancer. These features include cystic and solid tumour structure, the presence of septation and papillarities. Aletti et al. (2007) proposed that the presence of complex ovarian mass with both cystic and solid components and septation are highly suggestive of ovarian cancer. Other researchers (Twickler and Moschos, 2010) have found that the increased size of the mass is a significant factor in predicting malignancy. More recently, Valentin and colleagues confirmed in their study that unilocular adnexal cysts with papillation are more difficult to classify as benign or malignant using subjective assessment even when an experienced ultrasound examiner performs the scan (Valentin et al., 2013).

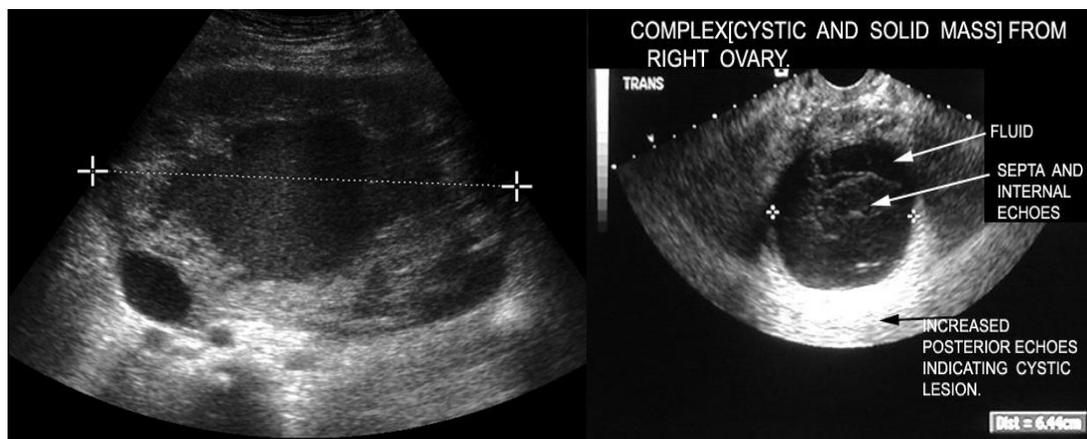


Figure 3: examples of B-mode images of ovarian tumours.

Although ultrasound findings are an excellent method for discriminating between benign and malignant adnexal tumours (Ameye et al., 2009), one of the main problems with this approach is that it is highly operator and equipment dependent and

therefore subjective (Wang et al., 2002, Shung, 2006). Therefore, the general concern in gynaecological ultrasound is the lack of standardized terms and procedures in the interpretations of the images (Timmerman, 2000). In order to provide a more standardized terminology and procedure, a group of investigators from the International Ovarian Tumour Analysis (IOTA) group has produced a paper entitled *“Terms, definitions and measurements to describe the sonographic features of adnexal tumours: a consensus opinion from the International Ovarian Tumour Group analysis (IOTA) group”* (Timmerman, 2000). Several scoring systems were developed to overcome the operator-dependent limitations and to increase the test performance, as well as allowing more dependable assessment of inter-observer and intra-observer variability and the comparison of results from different centres (Alcázar et al., 2003). According to Van Holsbeke et al. (2009), the main benefit of using scoring systems is to provide a helpful tool to allow less experienced ultrasound examiners to achieve the same diagnostic performance as an expert.

A study in China in 2011 showed that the operator’ level of confidence is positively associated with diagnostic performance. In addition, the conclusion was made that the accuracy of diagnosing ovarian tumours is greatly influenced by working experience and confidence score (Haiyan and Min, 2011).

Numerous studies have shown that the morphological appearance of adnexal masses on ultrasound can be used to predict malignancy (Timmerman et al., 2007, Varras, 2004, Berlanda et al., 2002, Morgante et al., 1999, Ferrazzi et al., 1997, Lerner et al., 1994). Both Lee and colleagues Lee et al. (2005) and Alcázar et al. (2003) reported that when using a morphological scoring systems, less experienced operators who

lack the expertise to make subjective assessments of adnexal masses would achieve a higher diagnostic accuracy that is comparable with the results of more experienced examiners.

A more recent prospective study (Alcazar et al., 2008) which included both pre- and post-menopausal asymptomatic women with adnexal masses found that an ultrasound scoring system provided good results with sensitivity and specificity of 95% in both of them, which was similar to previous reports (Guerriero et al., 2005, Berlanda et al., 2002), confirming that ultrasound-based triage is an excellent method of classifying adnexal masses. However, the scoring system misclassified adnexal masses when tumour size was > 10 cm, thus limiting its accuracy for large tumours. Alcàzar (2008) concluded that even in highly skilled hands, when using grey scale and Doppler ultrasound triage, a false-positive rate of 5-10% has to be anticipated in discriminating benign from malignant adnexal masses. Furthermore, the false-positive rate in using an ultrasound-based scoring system in the largest multicentre study was 24%. This was supported by Sokalaska et. al., which demonstrated that, even in the hands of an experienced examiner, it was not possible to make a conclusive diagnosis in adnexal pathology using subjective evaluation of grey-scale and Doppler ultrasound (Sokalska et al., 2009).

A study commenced in 2008 by Timmerman and colleagues (Timmerman et al., 2008) has been described as the largest study to analyse the ultrasound features of benign and malignant adnexal masses. They introduced a simple rule for describing adnexal masses, consisting of five ultrasound features of malignancy (M-features) and

five ultrasound features suggestive of benign mass (B-features). These features are presented in Table 1.

A mass is classified as malignant if at least one M-feature and none of the B-features are present, and vice versa. If no B- or M-features are present, or if both B- and M-features are present, then the rules are considered inconclusive (unclassifiable mass) and a different diagnostic method should be used. They found that to accurately determine the nature of an adnexal mass, subjective impression of the ultrasound morphology can be used. They agreed with previous studies (Modesitt et al., 2003) in that unilocular adnexal cysts have a very low risk of malignancy and that any other morphological appearance was associated with an increased risk of malignancy. However, a significant limitation to this study was that this rule cannot be applied to nearly 25% of tumours, since not all masses will present features clearly predictive of benignity or malignancy. The rule worked well for advanced invasive malignancies, but not for stage I borderline tumours and stage I primary invasive malignancies (Timmerman et al., 2008). In other words, it was applicable to tumours that were easily classifiable using operator pattern recognition but less apparent in more difficult tumours. Therefore, their results were incomplete.

Later, in 2013, Alcázar and colleagues commenced an external validation of the IOTA simple rule in a prospective study for which ultrasound scanning was performed by non-expert examiners. They concluded that the IOTA simple rule are undeniably simple, easy to learn and user friendly; additionally, their data yielded acceptable

results in terms of specificity (97.5%) in the hands of non-expert examiners. However, there was a 12% false-positive rate, which is relatively high, where 4 of 33 (12%) malignant masses were misdiagnosed as benign. This could be concerning if the simple rules were applied to triage for conservative management (Alcázar et al., 2013b).

B-features	M-features
B1: Unilocular cyst	M1: Irregular solid tumour
B2: Presence of solid components with largest diameter <7 mm	M2: Presence of ascites
B3: Presence of acoustic shadows	M3: At least four papillary structures
B4: Smooth multilocular tumour, with largest diameter <100 mm	M4: Irregular multilocular solid tumour with largest diameter ≥ 100 mm
B5: No blood flow (colour score 1)	M5: Very strong blood flow (colour score 4)

Table 1: Ultrasound features used in the International Ovarian Tumour Analysis (IOTA) simple rules.

2.1.2.2.1.1. Sassone Scoring System (see Table 2)

Sassone and colleagues (Sassone et al., 1991) devised a scoring system to characterize ovarian lesions using traditional grey-scale transvaginal ultrasonography relating ovarian morphology directly to risk of malignancy. It assigned a numeric value to

specific morphologic criteria, including inner wall structure, wall thickness, septal structure and echogenicity. The total score varies between 4 and 15, with a threshold value of 9 used to distinguish benign from malignant tumours. The researchers applied this scoring system in 143 women undergoing surgery for clinically detected pelvic masses, 20 of whom had ovarian malignancy. Using a score of ≥ 9 as indicative of cancer, the Sassone scoring was able to distinguish benign from malignant lesions with a sensitivity of 100%, specificity of 83%, positive predictive value (PPV) of 37% and negative predictive value (NPV) of 100%.

Value	Inner Wall Structure	Wall Thickness (mm)	Septa (mm)	Echogenicity
1	Smooth	Thin ≤ 3 mm	No septa	Sonolucent
2	Irregular ≤ 3 mm	Thick > 3 mm	Thin ≤ 3 mm	Low echogenicity
3	Papillaries > 3 mm	Not applicable, mostly solid	Thick > 3 mm	Low echogenicity with echogenic core
4	Not applicable, mostly solid			Mixed echogenicity
5				High echogenicity

Table 2: The original Sassone scoring system for adnexal masses (Sassone et al., 1991)

Ferrazzi et al. (1997) agreed with this finding and added that it has the advantage of being more descriptive and helpful for routine examination. Sassone's model was described by Geomini et al. (2009) as the most frequently validated scoring system. Furthermore, it has been reported and validated in numerous recent studies (Rossi et al., 2011, Geomini et al., 2009, Tempe et al., 2006, Alcázar et al., 2003, Mol et al.,

2001, Buckshee et al., 1998). However, several other studies that used Sassone scoring were excluded from the current study because the articles were published in languages other than English.

2.1.2.2.1.2. DePriest Scoring System (Kentucky score, see table 3)

Another similar scoring method was the morphology index described by DePriest et al. (1993). This index was based on tumour volume, wall structure and septal structure. A five-point scale (0-4) was developed within each category for specific criteria with the total points per evaluation varying from 0 to 12. Using a morphology index score of ≥ 5 as indicative of malignancy. In this study, this system had sensitivity of 89%, specificity of 73%, PPV of 46%, and NPV of 96% in postmenopausal women. The most reliable morphologic criterion in differentiating ovarian cancer from benign lesions was the wall structure (Varras, 2004, Geomini et al., 2009).

Category	0	1	2	3	4
Volume	<10 cm ³	10-50 cm ³	>50-200 cm ³	>200-500 cm ³	>500 cm ³
Cyst Wall Structure	Smooth <3 mm Thickness	Smooth >3 mm Thickness	Papillary Projection <3 mm	Papillary Projection ≥ 3 mm	Predominantly Solid
Septa Structure	No Septal	Thin Septal <3 mm	Thick Septal 3-10 mm	Solid Area ≥ 10 mm	Predominantly Solid

Table 3: The DePriest scoring system (DePriest et al., 1993)

In 2003, Ueland and his colleagues (Ueland et al., 2003) demonstrated that De Priest's Kentucky scoring system was an accurate and inexpensive method for distinguishing benign from malignant tumours. They explained that its only limitation was that the index was subject to interobserver variation.

Similar to Sassone, many groups have validated and compared the Kentucky scoring system with other scoring systems (Lee et al., 2005, Varras, 2004, Ueland et al., 2003, Van Nagell Jr and Ueland, 1999, Ferrazzi et al., 1997).

Although the use of a scoring system helps to improve test performance, the existence of multiple scoring systems may pose problems in clinical practice, mainly because there are too many of them and some of the parameters to be considered are often very complex.

2.1.2.2.1.3. Risk of Malignancy Index (see equation 2)

To improve the preoperative assessment of adnexal masses, most of these parameters (menopausal status, ultrasound findings and serum CA125) have been combined into diagnostic models. Such models have been developed to guide clinicians to identify patients with pelvic masses who are likely to have ovarian cancer. Risk of Malignancy Index (RMI) is a useful way of triaging women into low, moderate and high risk of malignancy (RCOG, 2003), which has been evaluated in numerous primary studies. Jacobs and colleagues developed the Risk of Malignancy Index (RMI) for referral of relevant women to gynaecologic oncologic centres (Jacobs et al., 1990) based on transabdominal ultrasound scan. Since then, various versions of the Risk of Malignancy Index have been published. However, the most commonly used version

in the UK is that published by the royal college of obstetricians and gynaecologists in October 2003 (RCOG, 2003)

$$\mathbf{RMI} = \mathbf{U} \times \mathbf{M} \times \mathbf{CA125}$$

Equation 2: RMI model

Where, U is an ultrasound score, and M is menopausal status, and the serum level of CA125 was applied directly to the formula.

Each of the following grey-scale morphological features was given one point when present: bilateral lesions, multilocular lesions, solid areas, intra-abdominal metastases and ascites. If the sum of these points was 0, an ultrasound score U=0 was given, while a sum of 1 point an ultrasound score U=1, and a score of U=3 was given when the sum of ultrasound points ≥ 2 (Jacobs et al., 1990, Yazbek et al., 2006).

RMI < 25	RMI between 25 - 250	RMI > 250
Low Risk < 3%	Moderate Risk 20%	High Risk 75%

Table 4: RMI categories as illustrated by (Prys Davies et al., 1993)

M is menopausal status, where Premenopausal status yielded M = 1 and postmenopausal status yielded M = 3.

The RMI was the first diagnostic model that combined demographics, sonographic and biochemical data in the assessment of women with adnexal masses. It provides a quantitative assessment of the risk of malignancy and can be used to differentiate

between benign and malignant lesions. At a cut-off level of 200, the sensitivity is 85% and the specificity is 97%.

The main advantages of the RMI in comparison with others are that it is a simple scoring system and a reliable, cheap, convenient and cost-effective method of preoperative discrimination between benign and malignant adnexal masses (van den Akker et al., 2010, Harry et al., 2009, Chia et al., 2008, Bailey et al., 2006, Andersen et al., 2003, Tingulstad et al., 1999, Jacobs et al., 1990).

It is worth noting that one of the issues of the RMI is that it is dependent on serum CA125, which is a non-specific marker for ovarian cancer: hence, it can lead to unnecessary surgical intervention (Ortashi, 2008).

Yazbek and colleagues (Yazbek et al., 2006) used the same cut-off value in their prospective observational study. In addition, they combined the RMI with another test called the Ovarian Crescent Sign (OCS) to diagnose ovarian malignancy. They concluded that both RMI and OCS are useful tests to discriminate between invasive and non-invasive ovarian tumours.

Tingulstad (1996) introduced RMI II by modifying the RMI scoring system described by Jacobs and re-evaluated its ability to differentiate benign from malignant ovarian lesions by utilizing Transvaginal instead of Transabdominal ultrasound, adjusting the ultrasound score to U=1 if no or one abnormality was seen rather than U= 0 for no abnormality. The reason for this was that Jacobs did not consider the value of serum CA125 level or the menopausal status in the case of no abnormality detected on

ultrasound with $U = 0$, will result in $RMI = 0$, regardless of the other parameters (Tingulstad et al., 1996).

Both RMI I and RMI II have been studied extensively in recent years and have been validated retrospectively and prospectively in numerous clinical studies where a cut-off value of 200 showed the best discrimination between benign and malignant adnexal masses, with high sensitivity 51-90% and specificity 51-97% levels (van den Akker et al., 2010, Lou et al., 2010, Geomini et al., 2009, Enakpene et al., 2009, Moolthiya et al., 2009, Harry et al., 2009, Ulusoy et al., 2007, Manjunath et al., 2001, Morgante et al., 1999).

Another study compared the accuracy of RMI I and RMI II, finding that RMI II is more sensitive than RMI I with higher specificity of 89% to 92%. In addition, it recommended the use of the RMI II scoring system due to its simplicity and reproducibility (Le et al., 2009).

Although two recent studies (Akdeniz et al., 2009, Moolthiya et al., 2009) agreed with the previous ones, they were excluded from this project because of the misuse of RMI I and RMI II. In the first study, the author applied an incorrect scoring for calculating the RMI, as he gave a score of $U = 1$ if no or one ultrasound feature was detected and a score of $U = 3$ if two or more of these features were detected, as well as giving a score of 2 for postmenopausal status instead of 3, which does not correspond to any of the RMI versions. The other study was excluded because the authors were trying to study the difference between RMI I and RMI II, except that they failed to identify the RMI

II correctly; they gave a score of $U=0$ for no features detected on ultrasound instead of $U=1$.

In 1999, Tingulstad further modified the RMI II by combining the ultrasound score of zero or one to give $U=1$, whereas for two or more features, $U=3$ was used in the equation. This was described as RMI III, and had a sensitivity of 78% and specificity of 92%. The authors concluded that RMI III was the ideal scoring system for referral of women with suspected malignant pelvic masses (Tingulstad et al., 1999).

A more recent study (Kader Ali Mohan et al., 2010), undertaken in an Australian population, compared the performance of all three versions of RMI (I, II and III) and reported that the best results were obtained with RMI II, which had 79% sensitivity and 88% specificity, compared to 71% sensitivity and 89% specificity for RMI I and RMI III, respectively. The authors also concluded that no statistical differences were observed in sensitivity and specificity values between the three versions of RMI. Similar results were found in a previous study by (Clarke et al., 2009).

Another study (Yamamoto et al., 2009) described a fourth RMI model (RMI 4) with the addition of tumour size score (S) as a fourth parameter. $RMI\ 4 = U \times M \times S \times CA125$, where a total ultrasound score of 0 or 1 yielded $U = 1$, and a score of 2 yielded $U = 4$. Premenopausal status yielded $M = 1$ and postmenopausal status yielded $M = 4$. A tumour size (single greatest diameter) of < 7 cm yielded $S = 1$, and ≥ 7 cm yielded $S = 2$. The authors compared the performance of RMI 4 with the previous three versions and concluded that RMI 4 was more reliable in discriminating malignant tumours from benign than RMI I, II or III when a cut-off value of 450 was

used. Nevertheless, due to the retrospective nature of the study, RMI 4 requires further validation with a prospective study (Yamamoto et al., 2009).

It has been recommended that a cut-off value of 200 is to be used for the RMI I in District General hospitals and health centres to increase referral to tertiary centres, and a cut-off value of 250 in specialized centres to maximize specificity and decrease false positives, thus reducing the number of interventions (Enakpene et al., 2009, Chia et al., 2008).

It is documented that scoring systems are beneficial in estimating the risk of malignancy in adnexal mass; however, they perform less well when they are used by inexperienced examiners compared to experts, as concluded by Van Holsbeke et al. (2009).

In 2011, an evidence review was published for NICE (National Institute for Health and Care Excellence) by the National Collaboration Centre for Cancer with the title of '*Ovarian Cancer: the Recognition and Initial Management of Ovarian Cancer*'. The recommendation was made in this review to use the RMI I as a malignancy index for women with suspected ovarian cancer based on the findings of the recent systematic review of diagnostic studies, which indicated that RMI I was superior in term of sensitivity and specificity to other comparators (Geomini et al., 2009) (NICE, 2011).

To summarise, scoring systems still suffer limitation of subjectivity. This is supported by the finding of the recent IOTA study, which recommended that simple rule should be used as the principal test to characterize masses as benign or malignant in

premenopausal women because they perform superiorly to RMI in this particular group (Kaijser et al., 2013).

2.1.2.2.1.4. PMI (Pelvic Mass Index)

This scoring system combines transvaginal ultrasonography with Doppler; it is independent of CA125. PMI assesses grey scale features such as size, laterality, presence of solid elements, septae and free fluid, all scoring one point each. The presence or absence of positive blood flow on Doppler ultrasound within the septa and/or solid component scores 2 points or -2 accordingly. Peripheral blood flow within ovarian stroma is not considered significant. The maximum score is 7 and the minimum is -2. Scores between -2 and 0 are considered low risk, scores between 1 and 2 intermediate and scores of greater than 3 are associated with high risk of malignancy. Sinha A. et al. (2015) aimed to validate the accuracy of PMI by comparing it to the RMI in a large patient cohort. They used 1,486 patients in a retrospective study over a seven-year period in a dedicated pelvic masses clinic in Wales. The authors explained that PMI is a fast way of estimating the risk of malignancy that leads to an accurate and reproducible diagnosis of the mass. They concluded that PMI is a useful tool in triaging patients with suspicious ovarian masses due to its high sensitivity (90.4%) and NPV (96.9%). When compared to RMI, it significantly outperforms RMI in diagnosing malignancy (AUC 0.823 vs. 0.770 respectively): (Sinha A et al., 2015)

2.1.2.2.1.5. ADNEX model

Recently the IOTA group developed a new model called the ADNEX model (the Assessment of Different NEoplasias in the adneXa). This model contains three

clinical and six ultrasound predictors: age, serum CA125 level, type of centre (oncology centres vs. other hospitals), maximum diameter of lesion, proportion of solid tissue, more than ten cyst locules, number of papillary projections, acoustic shadows, and ascites. Their aim was to develop a risk prediction model to preoperatively distinguish between benign, borderline, stage 1 invasive, stage 2 to 4 invasive, and secondary metastatic ovarian tumours. This huge study was performed in twenty-four ultrasound centres in ten different countries with a total of 5,909 patients (Van Calster et al., 2014).

Their final ADNEX model is available online at (www.iotagroup.org/adnexmodel/). This application has the advantage of calculating the risk even if the serum CA125 level information is unavailable, despite the decrease in performance. It was found that the ADNEX model has the potential to optimize management of women with adnexal masses in addition to offering excellent discrimination between benign and malignant masses, as alleged by (Van Calster et al., 2014). Moreover, the study showed that the proportion of solid tissue and serum CA125 level were the strongest predictors, while the type of centre was the weakest predictor, indicating that other predictors were determining the malignancy rate. However, in this study, the histology results were obtained in all masses because the model was based on patients who were selected for surgery, and therefore, the test performance could decrease if applied to tumours which were managed expectantly.

2.1.2.2.2. Three-dimensional ultrasound

Three-dimensional (3D) ultrasound was first used in 1989. In gynaecological imaging, the advantage of 3D over 2D ultrasound is that it improves the ability to visualize complex 3D structures (Prager et al., 2010).

According to (Kurjak et al., 2000a), three-dimensional display allows the operator to visualize many overlapping vessels easily and quickly as well as to assess their relationship to other vessels or surrounding tissues. It permits viewing of the structures in three dimensions interactively, rather than having to assemble the sectional images in the operator's mind.

(Kurjak et al., 2003) alleged that 3D sonography, when combined with 3D power Doppler, can significantly improve diagnostic accuracy in the assessment of suspected ovarian lesions, and supported their conclusion by citing previous studies demonstrating that 3D volume acquisition allows for careful evaluation of the internal surfaces of the cyst walls for outgrowths not seen by 2D technology (Bonilla-Musoles et al., 1995, Merz, 1999).

Another study that concurred with these findings found that evaluation with 3D ultrasound and 3D power Doppler can improve the diagnostic accuracy of ovarian tumours, where they had sensitivity of 90%, specificity of 89% and accuracy of 88% for prediction of ovarian malignancy when using 3D ultrasound (Laban et al., 2007).

In contrast, in the most recent review of three-dimensional ultrasound in gynaecological clinical practice, Alcazar disclosed that this technique is affected by some factors such as machine settings and attenuation, where there is a needs for

standardization. Therefore, it cannot yet be introduced into general practice. In addition, the author claims that there is a lack of robust data to support routine use of 3D ultrasound (Alcázar et al., 2012).

According to Wu and his colleagues, in their recent review to evaluate the present status and development of 3D ultrasonography in imaging the ovary, the three dimensions imaging improve spatial orientation and access multiplanar views, which provide additional information that can detect lesions not easily diagnosed by two-dimensional ultrasound. Moreover, they added that 3D ultrasound is a reproducible technique that improves the diagnostic accuracy for assessing ovarian cancer. However, it still needs more experience in training and operating than 2D ultrasound (Wu et al., 2012).

The role of 3D ultrasound in adnexal pathology is controversial and further research is needed in these areas to explore more potential uses of 3D ultrasound, as asserted by (Alcázar et al., 2012) and (Wu et al., 2012).

2.1.2.2.3. Contrast Ultrasound

In 2009, a study was conducted to examine the efficiency of intravenous contrast ultrasound called SonoVue to discriminate between benign and malignant adnexal masses. The authors concluded that ultrasound contrast examination is not superior to B-mode ultrasound techniques. Although SonoVue is a safe drug, it is rather expensive and the technique involves an intravenous injection. Furthermore, the acquisition of the information from the ultrasound image is difficult and the analysis is time-consuming, as asserted by the researchers, which makes this technique

impractical to use in differentiating benign from malignant masses (Testa et al., 2009).

2.1.2.2.4. Doppler

The Doppler effect enables ultrasound to be used to detect the motion of blood (Horskin, 2010). Doppler ultrasound has been used in medicine for almost forty years (Boote, 2003). It is used to measure blood velocity by means of Doppler frequency shift of the echoes received from red blood cells and allows the assessment of tumour vascularity (Rubin, 1994). The concept of Doppler is that malignant neoplasms have active blood vessel creation (angiogenesis) compared to normal or benign neoplasms. Benign lesions tend to form new tumour blood vessels peripherally from pre-existing blood vessels, whereas malignant tumours tend to form new tumour blood vessels centrally, as explained by (Jeong et al., 2000).

There is conflicting evidence as to whether adding colour Doppler imaging to ultrasound screening can reduce the rate of false positive test results or not (Tate et al., 2010). This is due to two issues that arise from colour Doppler ultrasound: first, the assessments are subjective, and second, the assessments depend on the quality of the equipment and the settings used (Timmerman, 2000).

Studies by Guerriero (2001, 1998) focused on the benefits of colour and power Doppler imaging to diagnose ovarian cancer. The earlier study explained that malignancy is suspected by power Doppler when arterial flow is visualized in an echogenic portion of a mass, unlike benign masses that have no similar arterial flow or when flow is seen only at the wall of the mass. It was concluded from that study

that power Doppler is helpful when B-mode is indecisive and that it could reduce the number of false positives and thus increase the diagnostic accuracy in atypical cases.

In the later study, it was recommended that at least one of the two Doppler techniques (conventional or power) should be used in conjunction with B-mode imaging as a secondary test (Guerriero et al., 2001, Guerriero et al., 1998).

A later study by (Tempe et al., 2006) examined the usefulness of colour Doppler in the preoperative assessment of ovarian tumours. They concluded that the overall effectiveness in diagnosing the type of lesion is enhanced when adding colour Doppler to ovarian morphology data.

All the studies on colour Doppler imaging revealed a significant overlap in Doppler flow indices between benign and malignant ovarian tumours (Alcázar et al., 2003, Ueland et al., 2003, Van Nagell Jr and Ueland, 1999, Guerriero et al., 1998). Furthermore, Ueland (2003) concluded that the addition of Doppler flow studies did not improve the diagnostic accuracy of the morphologic index.

In 2007, a book was published with a chapter titled: Ultrasound in ovarian carcinoma, in which the author discussed the performance of ultrasound in detecting malignancies based on morphological features and concluded that combining morphological and Doppler ultrasound assessment produces an ideal first imaging test for possible ovarian malignancies. However, the lower specificity of ultrasound requires further imaging evaluation, such as MRI, in patients where ultrasound is inconclusive (Webb, 2007).

2.1.2.2.4.1. Doppler indices

In Doppler arterial resistance techniques (Doppler signal analysis), a threshold value is used to characterise the mass. Parameters such as Pulsatility index (PI), resistive index (RI), and peak systolic velocity (PSV) have been used. A recent systematic review of the accuracy of ultrasonography with colour Doppler in ovarian tumours (Medeiros et al., 2009), showed that Doppler can detect malignancy or borderline lesions when RI is below 0.5 with a sensitivity of 87% and specificity of 90%. The authors concluded that Doppler is a useful preoperative test for predicting the diagnosis of pelvic masses and supported an earlier study (Kurjak and Predanic, 1993), which reported that the presence of vessels in the central, septal, or papillary projections, in conjunction with a diffuse vascular arrangement, and RI of less than 0.4, indicated that the mass was expected to be malignant. A possible limitation of this systematic review is the potential bias could be claimed due to all trials included were retrospective and there was a lack of blinding in their assessment.

These results were in agreement with another study (Erdogan et al., 2005) using Doppler ultrasound assessment in the diagnosis of ovarian tumours which showed that detection of Doppler signals in a solid component offered a precise preoperative method to differentiate between benign and malignant ovarian masses.

The major limitation of RI, PI and PSV is that the range of observed measurements in malignant masses overlaps with that observed in benign masses (Kurjak et al., 2003, Myers ER. et al., 2006). Therefore, Valentin et al. (1994) described Doppler as an impractical approach for diagnosing ovarian cancer from a clinical point of view. This argument is supported by another study (Laban et al., 2007), which concluded that RI measurements cannot be used alone for the detection of malignant ovarian tumours.

Furthermore, the authors explained that the reason for this is that there is considerable overlap between the RI measurements of benign and malignant ovarian masses. In addition, the overlap in RI range value between benign and malignant masses limits the efficiency of the application of threshold values: thus, cut-off values are not used.

Jeong et al. (2000) explained in their study that a comparison of different studies shows that no standard has been set concerning which Doppler index to use or what cut-off value is most appropriate. However, they found from previous literature that resistive indexes (RI) less than 0.4 and pulsatility indexes (PI) less than 1.0 are generally considered to be suspicious for malignancy. Additionally, the authors disclosed the problems that are associated with Doppler ultrasound, which include operator dependence and lack of standard criteria in distinguishing benign from malignant waveforms. Moreover, certain Doppler indexes can be misleading in premenopausal women due to physiologic alteration in the ovary during the menstrual cycle that cause lowered blood vessel resistance, thereby mimicking malignancy.

2.1.2.2.4.2. Three-dimensional Doppler

A new technique of Doppler ultrasound provides three-dimensional (3D) imaging. Three-dimensional ultrasound was first used in 1989 (Prager et al., 2010). Three-dimensional ultrasound utilizes the real-time capability of ultrasound to build a volume that can be constructed using high-performance work stations. (Hamid et al., 2011, Nelson, 2006).

To date, not enough information has been presented to determine whether 3D imaging of the adnexa adds significant information that is not available from conventional two-dimensional (2D) scanning (Benacerraf, 2008).

There are studies reporting that 3D power Doppler ultrasound may be useful for distinguishing benign from malignant ovarian tumours (Kurjak et al., 2003, Alcázar et al., 2005, Testa et al., 2005, Alcázar and Castillo, 2005). For example, Kurjak and colleagues claimed in their study that 3D power Doppler, when combined with the use of 3D sonography, will significantly improve the diagnostic accuracy of detecting stage 1 ovarian cancer, and supported their conclusion by citing evidence from a previous study conducted by (Cohen et al., 2001), which reported that 3D power Doppler better defines the morphological and vascular characteristics of ovarian lesions. Another study supported these findings and disclosed that the accuracy of diagnosing suspected ovarian lesions is significantly enhanced when using 3D ultrasound in combination with 3D power Doppler, and that this approach provides better visualization of tumour vascularity and could significantly improve the diagnostic accuracy in preoperative sonographic assessment in suspected ovarian lesions (Laban et al., 2007).

Alcazar studied tumour vascularity using 3D power Doppler in the early and advanced stages of ovarian cancer and found in his preliminary results that vascularization is higher in advanced stage and metastatic ovarian cancer than in the early stages (Alcázar, 2006).

In a previous group of studies in 1999, the researcher and his colleagues found that 3D power Doppler imaging can detect structural abnormalities of malignant tumour vessels, such as arteriovenous shunts. Therefore, it improves and facilitates the morphological and functional evaluation of benign as well as malignant pelvic tumours (Kurjak and Kupešić, 1999, Kurjak et al., 2000a, Kurjak et al., 2000b). Later, the same group of researchers demonstrated the ability of 3D Doppler ultrasound to perform as a secondary test in screening for ovarian cancer and described it as a novel approach for early detection of ovarian cancer (Kurjak et al., 2005).

Other studies suggested the need for further research (Alcázar, 2006, Rieck et al., 2006) and that it should be used as an adjunct to morphologic assessment (Wilson et al., 2006). Although Fishman et. al. (2001) stated that the clinical value of 3D ultrasound is promising for early detection of ovarian carcinoma and need to be investigated deeply, Dai et al. (2008) concluded that it did not improve the diagnostic accuracy for the prediction of malignancy in adnexal masses. Moreover, they further highlighted that 2D transvaginal sonography may still remain an important modality for the prediction of adnexal malignancy. In a preliminary study on 3D analysis of the vascularization of solid masses in the adnexal area, the authors justified the purpose of the study by citing the inaccuracy of 2D B-mode and colour/power Doppler to differentiate between benign and malignant tumours. They concluded that 3D quantitative analysis did not significantly improve the accuracy appreciated by 2D Doppler imaging (Testa et al., 2005).

In addition, Jokubkiene et al. (2007) found that objective quantification of colour signals of the tumour using 3D ultrasound did not appear to add more to B-mode

imaging when compared to subjective quantification using 2D power Doppler ultrasound. In a recent review, Alcázar and Jurado (2011) determined that additional studies are necessary to establish the role of 3D ultrasound in clinical practice in gynaecological oncology.

2.1.2.3. CT imaging

CT has several advantages: it is widely available and can be done rapidly and easily. Moreover, CT of the abdomen or pelvis allows comprehensive evaluation of all possible sites of the primary tumour. This modality has a major advantage over US and MRI imaging, as it allows oral contrast agents to mark the bowel and help differentiate bowel from peritoneal implants. Therefore, CT is a very attractive method for evaluating the spread of the disease in women with adnexal malignancy. Nevertheless, available studies have not proven that CT is significantly superior to other modalities in staging ovarian malignancy.

Jeong et al. (2000) pointed out that many studies have revealed that CT is neither sensitive enough nor specific enough to replace laparotomy and that the largest study to date comparing US, CT and MRI in the staging of ovarian cancer demonstrated little difference between the modalities (Kurtz et al., 1999).

In contrast, a recent study demonstrated a higher sensitivity and specificity (79.2%, 91.6%) when compared to transvaginal ultrasound (51.9%, 87.9%) respectively (Firoozabadi et al., 2011). However, this study has some pitfalls: for example, the examination and interpretation of the scans were carried out by different people, but their level of experience was not mentioned, and the sample size was relatively small (139 patients) compared to other similar studies.

Another study described CT and MRI as complementary imaging techniques and further explained that they can be used as an adjunct to ultrasound in specific cases but cannot be used as a first imaging modality of choice in tumour staging (Fischerova, 2011).

2.1.2.4. Magnetic Resonance Imaging (MRI)

The principle advantage of MR imaging is that it combines some of the best features of CT and US. Numerous types of tissue and fluid can be discriminated using MR imaging using their signal intensity characteristics. Jeong and his colleagues stated that malignant tumours do not have specific MR imaging signal intensity characteristics, so these tumours must be distinguished based on morphological criteria (Jeong et al., 2000).

In most studies, MR imaging has proved superior and more accurate than endovaginal US in differentiation between benign and malignant adnexal masses; however, it is more expensive, time consuming and impractical to perform MR imaging in all patients with abnormalities (Yamashita et al., 1995, Komatsu et al., 1996, Yamashita et al., 1997).

In 1999, Kurtz and colleagues performed a comparative study to determine the optimal imaging modality for diagnosing and staging ovarian cancer. Their results showed a noticeable accuracy in demonstrating the extent of the malignant spread for all the three imaging modalities (US, CT and MRI) and they all had high staging accuracies at ROC curve analysis of 0.91. However, they further explained that MRI

is superior to Doppler and CT in accurately diagnosing complex or solid ovarian masses at stage 3 malignant lesions (Kurtz et al., 1999).

In a prospective study of women with suspected adnexal masses, both Doppler ultrasound and MRI were highly sensitive for identifying malignant lesions (ultrasound 100%, MRI 96.6%); however, the specificity of MRI was significantly greater (ultrasound 39.5%, MRI 83.7%): therefore, women who clinically have a low risk of malignancy but who have complex sonographic morphology may benefit from MRI (Sohaib et al., 2005).

Similarly, a more recent study by Iyer and Lee described the role of MR, CT and PET/CT in detecting ovarian cancer. They found that MR imaging is useful as a secondary imaging technique for further investigating masses to define the extent of the disease when ultrasound is indecisive. Lesions that are indeterminate on ultrasound can often be diagnosed with high specificity by MRI (Iyer and Lee, 2010).

A recent systematic review and meta-analysis of the preoperative identification of suspicious adnexal masses by Dodge and his colleague is presented in Table 5. The results showed that 3D ultrasound has higher sensitivity and specificity when compared to 2D ultrasound. Furthermore, morphological scoring systems revealed a respectable performance in their sensitivity and specificity. In contrast, colour Doppler was neither sensitive nor specific in the assessment of adnexal masses when compared to ultrasound. The combination of morphology and Doppler assessment had a higher sensitivity and specificity than either modality alone. In addition, of the three imaging modalities considered, MRI appeared to perform the best, even though the results were not statistically different from CT. Finally, the measurement of the

CA125 tumour marker seems to be less reliable than other available preoperative methods (Dodge et al., 2012).

Method		Pooled Sensitivity	Pooled Specificity
Sassone scoring system (cut-off point of 9)		88.6%	77.5%
Lerner scoring system		90%	63%
DePriest scoring system		91%	69%
Risk of malignancy index (RMI)	RMI I (Cut off of 200)	79.2%	91.7%
	RMI II (Cut off of 200)	79%	81%
	RMI III (Cut off of 200)	74%	91%
Doppler sonography	2D power Doppler	49-100%	74-100%
	3D power Doppler	68%-100%	40-98%
	Resistance index (RI)	77.2%	89.8%
	Pulsatility index (PI)	80.6%	79.9%
Combined morphology and Doppler	2D ultrasound plus Doppler	91%	91.7%
	3D ultrasound plus Doppler	97.8- 100%	79.2- 84.2%
Other imaging modalities	MRI	91.9%	88.4%
	CT	87.2%	84%
CA125 Cut off of 35 U/ml		78.7%	77.9%

Table 5: Summary of the systemic review of preoperative methods to diagnose adnexal masses (Dodge et al., 2012).

Moreover, the most recent systematic review and meta-analysis, performed by Kaijser and his colleagues, which focused on the presurgical diagnosis of adnexal masses using mathematical models and scoring systems, update the results of previous systematic review (Dodge et al., 2012). The findings of this study are summarised in

Table 6. This update of the previous systematic review and meta-analysis demonstrates that the IOTA simple rule is currently the best diagnostic test available for diagnosing ovarian masses in premenopausal women (Kaijser et al., 2014).

Model	Cut-off	Sensitivity	Specificity
Sassone	>9	85%	80%
Lerner	>3	80%	61%
Depriest	>5	90%	68%
Simple rule	n/a	93%	81%
RMI I	200	72%	92%
RMI II	200	75%	87%
RMI III	200	70%	91%
RMI IV	450	68%	94%

Table 6: Pooled summary estimates of sensitivity and specificity in the new systematic review (Kaijser et al., 2014)

2.1.2.5. Laparoscopy and Laparotomy

Laparoscopy is the insertion of a thin lighted tube (called a laparoscope) through the abdominal wall in order to inspect the inside of the abdomen and to remove tissue samples, while laparotomy is a surgical incision made in the wall of the abdomen (Myers ER. et al., 2006). The decision for surgery depends on the probability of malignancy (Kinkel et al., 2000). Furthermore, laparotomy is performed if there is a strong suspicion of malignancy (Benedet et al., 2000), while laparoscopy is frequently used in the diagnostic evaluation of adnexal masses. However, laparoscopic diagnosis raises concerns about the possibility of tumour spillage due to cyst rupture, which correlates with a worse prognosis. In addition, laparoscopy may lead to the diagnosis and resection of large numbers of functional cysts and other lesions that could have been followed up clinically without surgery. For these reasons, laparoscopy is reserved for patients with masses that are non-suspicious based on imaging findings (Jeong et al., 2000).

In young women with non-malignant ovarian lesions such as endometriosis and benign cysts, treatment with laparoscopy can avoid laparotomy procedures (Jeong et al., 2000). Unlike young women, in postmenopausal women who are not suitable for conservative management, oophorectomy is recommended by RCOG even when the risk of malignancy is low (RCOG, 2003). According to Jeong et al. (2000), exploratory laparotomy is necessary in all cases of suspected ovarian cancer to confirm the diagnosis, determine the extent of the disease, and to resect the tumour.

Recently, a prospective observational study was commenced to assess long-term outcomes of expectant management for persistent adnexal masses in asymptomatic premenopausal women. They found that expectant management of cysts with benign ultrasound morphology is an option, especially as a significant proportion of masses resolved spontaneously with a very low risk of torsion (0.4%) and of cancer (0.9%: (Alcázar et al., 2013a).

2.1.2.6. Texture analysis

Image analysis methods are important to assist in detection and diagnosis. They have been developed to help physicians acquire diagnostic information and improve clinical decisions. Studies show that radiologists do not detect all abnormalities on images that are visible on retrospective review and do not always characterize the abnormality correctly. This can be caused by limitations in the human eye-brain visual system, reader fatigue, distraction, the presence of overlapping structures that camouflage disease in images, and the large number of normal cases seen in screening programs (Giger et al., 2008).

Therefore, researchers have started to investigate computerized image analysis that aims to automate the detection of abnormalities, including analysis of breast images and chest radiographs (Toriwaki et al., 1973). However, it has been difficult to achieve the accuracy and the acceptance required for clinical use, as asserted by Giger et al. (2008).

Later, in the mid-1980s, a team of researchers at the University of Chicago dedicated their efforts to computer-aided diagnosis (CAD), which means using computer output as an aid to radiologists rather than completely automatic computer interpretation (Chan et al., 1987, Giger et al., 1987). Furthermore, Giger et al. (2008) defined CAD as a diagnosis made by a radiologist who uses the output from a computer analysis of the image data in their decision-making process, so that the final medical decision is made by the radiologist, not the computer. They explained further that the role of the computer analysis is not to replace the radiologist, but rather to aid him/her in image interpretation and decision making.

Since then, the growth of CAD has been tremendous over the past twenty years. It has spread widely and quickly (Doi, 2007). Currently, CAD has been extended to include image analysis of various disease types, such as breast cancer, lung cancer, interstitial disease, colon cancer, osteoporosis, and vascular plaque aneurysms, using images acquired by different modalities, such as ultrasound, CT, PET, MRI and others.

The goal of CAD is to reduce search, interpretation errors, and variation between and within observers. In an ultrasonic image, different tissues always have significantly different textures.

Generally, texture of images refers to the appearance, structure and arrangement of the parts of an object within the image (Castellano et al., 2004). Texture cannot be precisely defined due to its wide variability (Mathias et al., 1999). In digital images, the concept of texture may be attributed to the distribution of grey-level values across

the pixels of a given region of interest in an image. In other words, texture analysis is a technique for evaluating the position and signal intensity characteristics of pixels. (Livens et al., 1997, page 581) revealed a more formal definition of texture as “*the set of local neighbourhood properties of the grey levels of an image region.*”

According to Srinivasan and Shobha, texture analysis refers to a class of mathematical procedures and models that characterize the spatial variations within imagery as a means of extracting information. They explained in their study that there is no single method of texture representation that is adequate for a variety of textures, since texture has so many different dimensions (Srinivasan and Shobha, 2008).

Texture analysis can be divided into categories such as structural, model-based, statistical, and transform, depending on the means utilized to evaluate the interrelationship of the pixels (Castellano et al., 2004, Holli et al., 2010). Please refer to Figure 4 for an illustration.

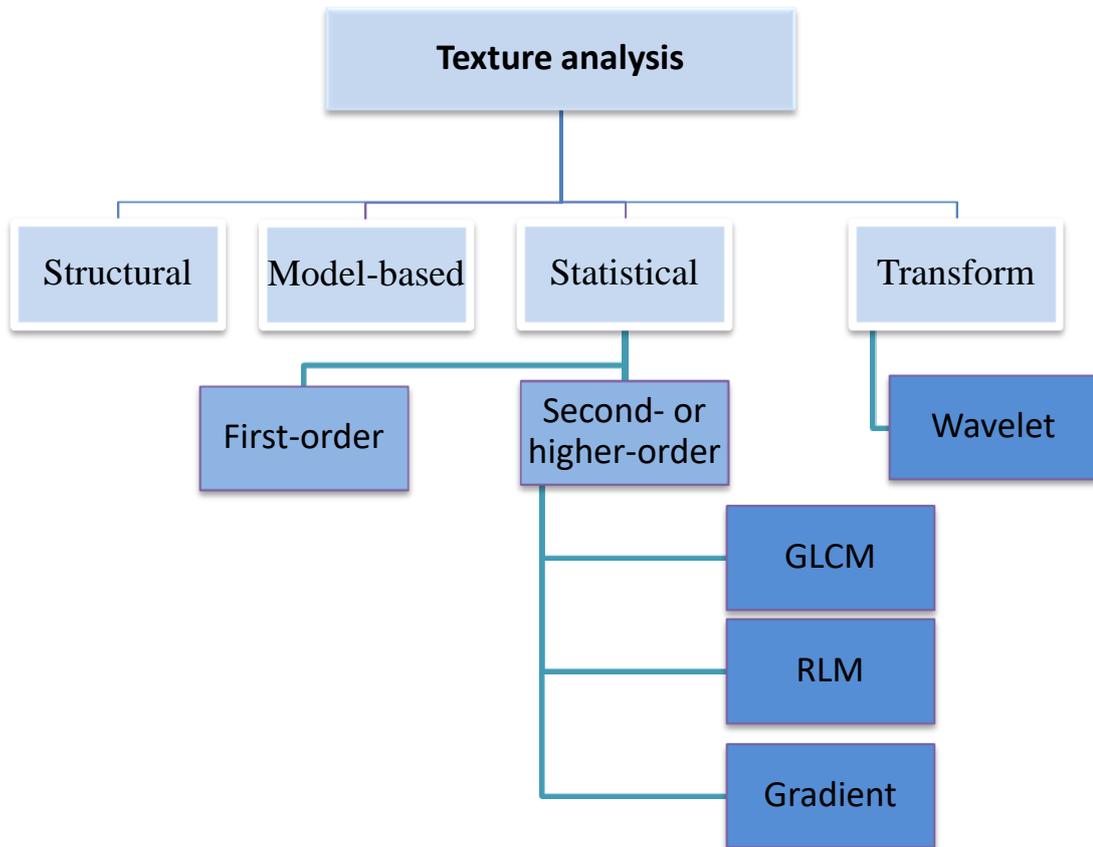


Figure 4: Types and sub-types of texture analysis.

In medical images, statistical methods are the most widely used. These methods analyse the spatial distribution of grey values by computing local features at each point in the image (Sirinivasan and Shobha, 2008). According to Tuceryan and Jain (1998), statistics are classified as a first-, second- or higher-order statistics according to the number of points which define the local feature. In first-order statistics, image properties depend merely on a singular pixel value, whereas second-order statistics are properties of pixel pairs.

First-order statistics do not consider pixel neighbourhood relationships. Common features include mean grey scale, standard deviation of the mean, skewness (deviation of the pixel distribution) and kurtosis (steepness of the pixel distribution), all of which can usually be detected visually. The limitation of this method is that it provides no information about the position of pixels relative to each other, due to texture analysis being based solely on the grey level of the histogram (Sirinivasan and Shobha, 2008).

The second order statistical method uses grey-level run-length measures and the grey-level co-occurrence matrix (GLCM). The latter shows how often each grey level occurs at a pixel located at a fixed geometric position relative to each other pixel, as a function of grey level. In this method, higher discrimination indices are obtained and cannot be visually detected. For that reason, second or higher order statistical methods are used in medical images texture analysis (Holli et al., 2010). The grey level co-occurrence matrix (GLCM) is a measurement of an image's statistical properties, visual characteristics, using information theory measure and correlation based information. (Xian, 2010)

The co-occurrence matrix is the joint probability occurrence of grey level of two pixels within a defined spatial relationship in an image (Sharma and Singh, 2001). This feature will be explained in more depth in Chapter 4.

2.1.2.6.1. Application of texture analysis to medical images

Numerous studies have demonstrated the value of texture analysis in the medical field. Image texture of medical images defines the internal structure of human tissue or organs (Szczypiński et al., 2009) as well as pathological changes (Xian, 2010).

Texture analysis techniques in medical images were first reported in the 1960s, and acknowledged by Doi (2007). They have been useful in various types of diseases, such as diseases of the liver (Lee et al., 2003), thyroid (Smutek et al., 2003), breast (Alacam et al., 2003, Ramos et al., 2012), kidney, pancreas (Das et al., 2008), heart (Tsai and Kojima, 2005) and coronary arteries (Nailon et al., 1996) as well as brain tumours (Herlidou-Même et al., 2003).

Texture analysis has also been applied to several imaging modalities such as ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), conventional x-ray and mammography. According to (Tsai and Kojima, 2005), the employment of texture analysis in medical imaging has proven to be valuable, especially for MRI, CT and ultrasound. Most of the previous works done in texture analysis encompass MRI images due to the great amount of details provided by this technique (Castellano et al., 2004). Texture analysis of all kinds of images is possible and has been achieved in literature.

2.1.2.6.1.1. Texture analysis of ultrasound images

Many researchers have focused their studies on quantifying the echo signal in B-mode scans. The principle of texture analysis in ultrasound images was explained by (Morris, 1988): if the structure of the tissue is affected by the disease process, this will

result in alteration of the ultrasound signal, meaning that the statistical measures will differ from those of normal tissue. For example, the transformation of cancerous tissue will change the tissue characteristics, such as density and elasticity. Therefore, textural features derived from cancerous tissue will differ from normal tissue based on the above-mentioned principle.

Nailon and Spencer investigated the use of the statistical texture analysis technique to assess intravascular ultrasound in order to characterize intracoronary thrombus. Results from their study were difficult to interpret due to their relatively small data set (Nailon et al., 1996).

Similarly, Vince and colleagues conducted a study that aimed to evaluate five texture analysis techniques and determine their ability to distinguish between plaque lesions of different compositions and found two of them to be adequate (Vince et al., 2000).

In 2003, Smutek and colleagues used texture analysis in ultrasound to diagnose chronic inflammation of the thyroid gland. Their results showed that it is possible to gain objective, quantitative characteristics of the thyroid gland and use it for tissue classification (Smutek et al., 2003).

Similarly, texture analysis was used to aid in the classification of breast tumours using ultrasound (Huang et al., 2008). Their results showed that analysis of sonographic characteristics can assist in differentiating between benign and malignant lesions. This study was in agreement with an early preliminary study done in 1988, which explained that texture analysis systems were applied successfully to tissue characterization in breast tumours. It concluded that applying texture analysis to a

binary classification of tumours into benign and positive classes was possible. Moreover, the application of texture analysis to the population studied reduced the incidence of false positive diagnosis by 40% when compared to cases of tumours that were readable by ultrasound. They also suggested that texture analysis is superior to professional radiologist working with the same images.

In another study conducted by Chen et al. (2002), wavelet transform was used to diagnose solid breast masses and this performed well for breast tumour diagnosis.

Texture analysis has also been studied extensively on liver-related diseases. A recent study by Xian (2010), using ultrasound, showed that the grey-level texture features technique was a feasible and excellent classification of liver tumours. Texture analysis can improve the diagnostic rule of the B-mode ultrasound images, as asserted by Vicas et al. (2011), who studied texture analysis as a non-invasive tool for the assessment of chronic hepatitis C. However, the role of human experts could not be reduced or eliminated in this study.

In a slightly different field of ultrasound, a study was conducted to evaluate some of the analysis techniques in monitoring follicular response in ovulation induction. It demonstrated that analysis of ultrasound images can aid in developing a reliable non-invasive technique to improve the clinical management of women undergoing ovarian stimulation (Pierson and Adams, 1995).

Texture analysis of ultrasound images has also been applied to prostate cancer. In 1993, Basset and colleagues used second-order statistics, namely co-occurrence matrices, to identify prostate tumours. Their results yielded fairly good tissue

signatures obtained with parameters derived from these matrices. In addition, 78% of the samples were classified with success, which was a high score considering that the images could not be discriminated visually (Basset et al., 1993).

Braeckman and colleagues studied computer-aided ultrasonography for detecting prostate cancer using an ultrasound-based technology called HistoScanning. This is a technology that detects specific changes in tissue morphology by extracting and quantifying statistical features from backscattered ultrasound data. Furthermore, it has been developed to distinguish cancerous and noncancerous tissues in solid organs. They concluded that HistoScanning is an accurate non-invasive method to detect cancer foci and suggested that it might be useful as a triage test for prostate cancer (Braeckman et al., 2008b, Braeckman et al., 2008a).

More recently, an attempt to create an online paradigm to characterise ovarian tumours was made by Acharya and colleagues. They used 3D transvaginal images in their analysis and applied several texture features including GLCM and Run Length Matrix (RLM). It was concluded that the use of CAD techniques using a combination of four texture features techniques has a good sensitivity of 94.3% and specificity of 99.7% and has the advantage of being objective. However, this system has been tested on only twenty cases and needs to be clinically validated to assess the diagnostic accuracy of this method (Acharya et al., 2013).

2.1.2.6.1.2. Texture analysis of CT images

Texture analysis has also proven to be beneficial in improving the interpretation of CT images. For example, Gletsos et al. (2003) studied the grey level co-occurrence

matrix technique as a method to classify four types of liver tissue: normal liver, hepatic cysts, haemangioma and hepatocellular carcinomas. Another study investigated the feasibility of the same technique on CT images of interstitial lung diseases and reported a sensitivity and specificity of 73-93% and 90-98% respectively (Xu et al., 2006).

Two most recent studies were initiated by the same group (Kumar and Moni): the first, in 2010, studied the characterization of liver tumours by CT and disclosed promising results, recommending the technique to be applied in the diagnosis of other types of liver disease. The second study, in 2012, proposed an automated computer-aided diagnosis (CAD) system for recognizing different types of liver tumours by analyzing tumour texture images using texture recognition techniques. They achieved 94% accuracy of the classifier when compared to other methods. In addition, they suggested that the proposed system can be extended to the diagnosis of other types of liver diseases as well (Kumar et al., 2012).

2.1.2.6.1.3. Texture analysis on MRI images

Similar to CT images, texture analysis has been also demonstrated to be useful in improving the interpretation of MRI images. It was first applied in the early 1980s (Herlidou-Même et al., 2003). Texture analysis has the advantage of discriminating complexity and offers information that is not visible to the human eye (Herlidou-Même et al., 2003).

Texture analysis on MRI images has been applied to multiple sclerosis related studies. In 1999, a study commenced using texture analysis to quantify the pathological

changes that occur within the spinal cord associated with multiple sclerosis. The spatial grey-level co-occurrence method was used in this study, justifying it from previous studies that applied this method in relevant application and because it performed well for small regions. Their results showed significant differences in texture between normal and multiple sclerosis images. Furthermore, texture is valuable in detecting changes in pathology early in the disease before spinal cord atrophy occurs (Mathias et al., 1999). Another recent study on multiple sclerosis achieved an accuracy of 88.4% (Theocharakis et al., 2009).

According to Harrison et al. (2010), when classification of white matter and multiple sclerosis lesions were studied, excellent distinction was achieved when using texture measures with an accuracy between 96% and 100%. In the same year, another study demonstrated significant changes in texture measures of cerebral tissue between hemispheres and corpus callosum segments in traumatic brain injury patients. It was suggested that this technique may be used as a novel additional tool for identifying the invisible changes in cerebral tissue in mild traumatic brain injury, aiding clinicians to make an early diagnosis (Holli et al., 2010).

2.1.2.6.1.4. Texture analysis on Mammography images

A recent study conducted by (Ramaraj and Raghavan, 2011), which focused on wavelet techniques for cancer diagnosis, found that the wavelet transform is an excellent tool to investigate mammograms, MRI and ultrasound breast images. These results are in agreement with a previous study in 2007 where a high accuracy rate was found (89%) when investigating the feasibility of texture analysis in differentiating benign from malignant breast tissue (Karahaliou et al., 2007).

Beside ultrasound, CT, MRI and mammography, texture analysis techniques have also been applied to other imaging modalities, such as colposcopic imaging (Ji et al., 2000) optical tomography (Baroni et al., 2007, Gossage et al., 2003) and capsule endoscopy (Li and Meng, 2009).

2.2. Summary

Ovarian cancer is the second most common gynaecological malignancy; however, it remains the leading cause of death among these diseases. In spite of the diagnostic and therapeutic advances in the care of women with ovarian cancer, the overall five-year survival rate remains unchanged. The reason for this is that most cases are diagnosed in the late stages of the disease, when the five-year survival rate falls below 20%.

Therefore, early detection of ovarian malignancy is of great clinical importance. Ultrasonography is currently considered the primary imaging modality for diagnosing adnexal masses. However, ultrasound is operator-dependent and thus the accuracy and reproducibility of the diagnosis are subject to the experience of the operator.

In order to reduce operator dependency, texture analysis, which is able to quantitatively characterize tissue through texture content, will be used in this study to objectively differentiate between normal, benign and malignant ovarian tissue.

Texture analysis is defined as the spatial distribution of the pixel grey value (intensity) of B-mode images. It is a descriptor of local brightness variation from pixel to pixel in a small neighbourhood through an image. Two features of texture analysis were tested in previous pilot research and suggested to be robust and demonstrated good repeatability in characterizing ovarian malignancy: these are grey level co-occurrence matrix (GLCM) and wavelet analysis. Therefore, they will be used in this study, along with applying another modality called Acoustic Structural Quantification (ASQ).

A previous pilot study suggested that texture analysis is beneficial in objectively differentiating ovarian lesions (Hamid, 2011). This study will involve a large number of patients and is expected to help confirm and validate the sensitivity and specificity of this method.

The texture analysis technique will be compared to widely-used scoring systems such as the Risk of Malignancy Index (RMI), the Pelvic Mass Index (PMI) and the ADNEX model, and the ASQ texture analysis technique, will also be applied.

3. Acoustic Structure Quantification (ASQ)

This chapter will introduce and test a new method called ASQ. It is divided into four main sections: section A will include a definition of the new term, followed by some background information and a phantom study where the repeatability and reproducibility will be tested, while section B will study the influence factors, such as ROI size, ROI depth, Focus, Gain setting and frequency, on ASQ. Then, in section C, the influence of pre-defined image parameters on ASQ output will be discussed. Lastly, section D will demonstrate the application of ASQ to images of benign and malignant masses.

3.1. (A) Introduction, how ASQ works and background

Ultrasonography represents an excellent examination modality that is non-invasive, inexpensive and can be performed repeatedly with no risk to the patient. However, the diagnostic performance of ultrasonography in general depends on the empirical and qualitative reading skills of the examiner (Kuroda et al., 2012). The human eye cannot distinguish between more than fifteen to twenty shades of grey in ultrasound images (Lagalla and Midiri, 1998): it is therefore worthwhile developing a new technique that can overcome these limitations.

Acoustic structure quantification (ASQ) is new software that analyses the statistical information of the acquired (receiving) echo signals (Toshiba Medical Systems, Europe). This is accomplished by looking at the speckle pattern in a certain region of interest (ROI) (De Kant, 2011). It is a non-invasive tool that can assist the ultrasound operator in the assessment, characterisation and follow-up of fibrotic disease during a standard ultrasound scan. Liver ASQ assesses tissue homogeneity quantitatively and depicts tissue properties in a convenient colour-coded display. Therefore, the aim of this chapter is to investigate the suitability of ASQ in diagnosing ovarian cancer.

3.1.1 How ASQ works

Although ASQ is ready-to-use software, it was thought that a brief explanation of how it works is necessary. It is a software that analyses the statistical information contained in the received echo signals: it takes the raw data in a ROI and then analyses the speckle pattern, which produces a parametric image and quantification of tissue type for tissue characterization. Some basic concepts of ultrasound have to be introduced to understand the ASQ concept. These include the origin of ultrasound speckle, the Probability Density Function (PDF), the χ^2 function, and the modified χ^2 function called C^2 . Having explained these concepts, the process of how ASQ works can be understood. First, the ultrasound speckle is a random, deterministic interference pattern in an image formed when ultrasound interacts with many sub-resolution scatterers. This speckle pattern is an interference pattern formed as a result of ultrasound scatter.

In general, for a simple reflection, the dimensions of the reflecting surface must be greater than several wavelengths of the ultrasound wave. When small targets, smaller than the wavelength of ultrasound, are scanned, it will cause scatter called *Rayleigh Scattering*. Please refer to Figure 5:

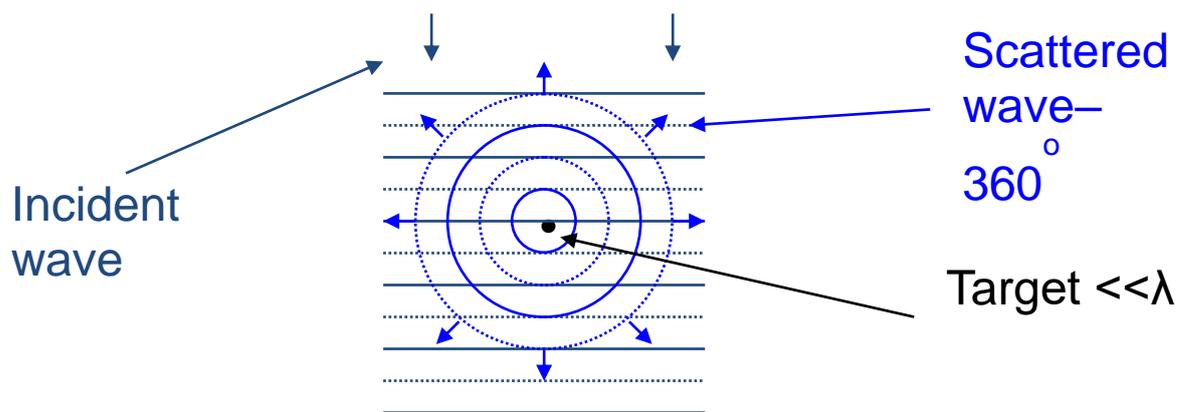


Figure 5: Rayleigh scattering (adopted from Toshiba catalogue)

Therefore, a tissue composed of structures $< \lambda$ of the ultrasound will produce a speckle pattern. The texture of the observed speckle pattern as a result of scattering does not correspond to the underlying structure, which means no spatial information. The local brightness (amplitude) of the speckle pattern, however, does reflect the local echogenicity of the underlying scatterers.

Because speckle is random, the echo signal from scatterers alone has a zero mean, two-dimensional Gaussian probability density function (PDF). This is the Rayleigh PDF, a function that describes the relative likelihood that this random variable will take on a given value (De Kant, 2011). Please see Figure 6.

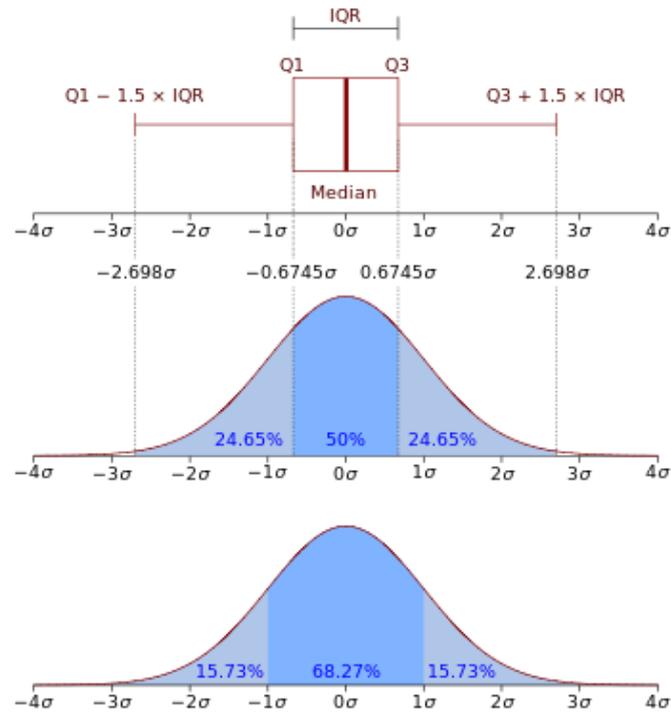


Figure 6: Rayleigh PDF

This means in practice that if tissue is made up of scatterers alone (no reflective surfaces), the resulting image is a representation of the PDF for that tissue.

Second, the χ^2 function is simply a statistical test which compares one distribution against another. Please refer to equation 3.

$$\chi^2 = (n - 1) \frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}$$

Where

$\hat{\sigma}^2$ = the variance of the sample (abnormal tissue)

$\hat{\sigma}_0^2$ = the variance of the population (normal tissue)

Equation 3: The χ^2 function

If the sample is exactly the same as the normal population, $\chi^2 = 1$.

Therefore, when ASQ is first applied, it uses a modified χ^2 called C^2 . Where C^2 is divided by degree of freedom (n-1) to get the averaging and normalize the test. Please see equation 4.

$$C^2 = \frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}$$

Equation 4: C^2 equation

Secondly, a population variance is required that corresponds to normal liver; however, it is difficult to achieve in a diseased liver, and it is also difficult to use a standard value from the normal livers of other patients, because the variance changes due to machine settings, overlying tissue etc.

So to overcome this problem, we calculate an approximate value from the average of samples taken in the user defined ROI by applying certain thresholds, then:

$$\hat{\sigma}_0^2 = \sigma_R^2(\mu_m)$$

Equation 5

Thresholds are also applied to calculate a modified $\hat{\sigma}^2$, so

$$\hat{\sigma}^2 = \sigma_m^2$$

Equation 6

Finally we get ASQ:

$$C_m^2 = \frac{\sigma_m^2}{\sigma_R^2(\mu_m)}$$

Equation 7: ASQ equation

To produce C_m^2 a large ROI is drawn, avoiding reflective surfaces. Many (~300) small secondary ROIs (about twice the size of the resolution of the system) are automatically created within the large ROI. This raw data is stored as a 2x2 matrix for and the mean and variance in each small ROI is then calculated. A histogram of C_m^2 to then produced.

If the tissue is completely homogenous, then $C_m^2 = 1$ and the histogram will show a normal distribution because the sample (numerator) and the normal tissue (denominator) will be the same. Please see Figure 7.

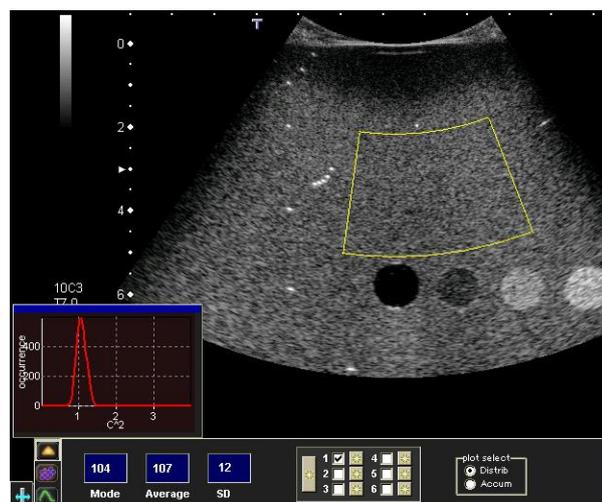


Figure 7: normal distribution in histogram of a homogenous phantom.

However, if the sample variance (numerator) is different due to heterogeneity in the sample, $C_m^2 > 1$ the bandwidth of the histogram will increase and the curve will be shifted to the right: please refer to Figure 8.

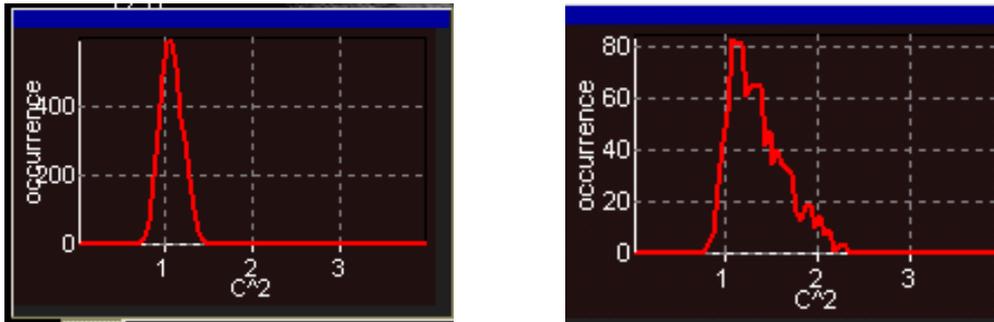


Figure 8: Example of homogenous tissue histogram and non-homogenous tissue.

If there is a point reflector in the image, it will show as discrete, high amplitude signals, which will affect the C_m^2 curve (produce a tail to the curve). This tail is removed from the analysis and displayed as a separate entity called the blue curve. Please refer to figures 9 and 10.

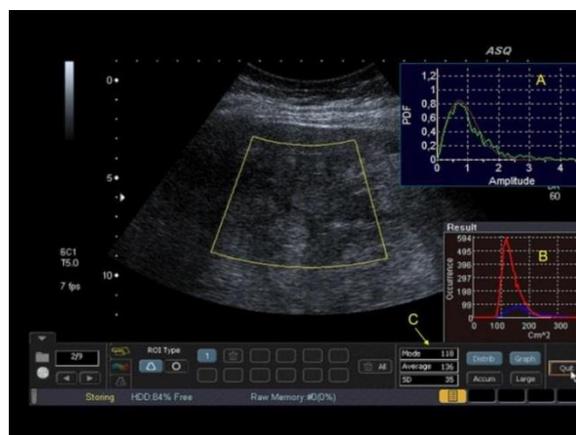


Figure 9: The tail in the curve produced by a discrete high amplitude signals.

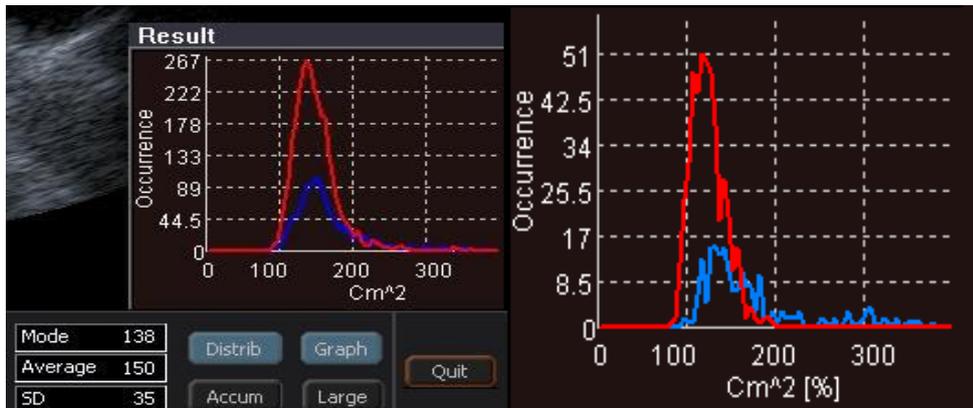


Figure 10: The blue curve in the histogram.

Additional analysis information produced in the ASQ software is the Q-Q graph. This graph shows the variation between the “normalised” PDF (calculated denominator) and a theoretical PDF. The more closely these lines are matched, the better the denominator value used in the C_m^2 (Please refer to Figure 11).

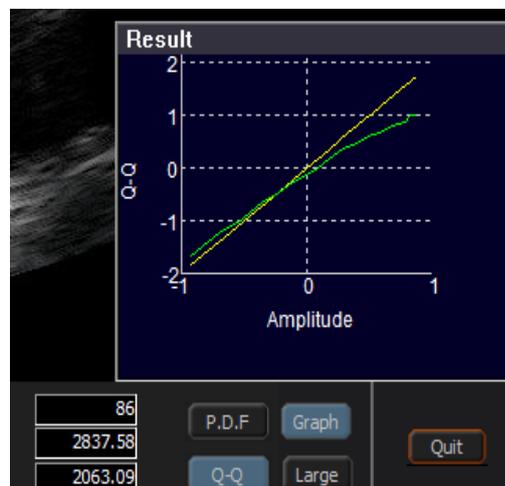


Figure 11: The Q-Q graph in ASQ window.

ASQ offers two modes of presentation: the first is a statistical graph, which draws distribution curve parameters in response to the fibre structure that reflects the ultrasound beam (shown in figure10). The second display is like familiar images with colour Doppler of the different values superimposed on a B-mode image (Hung, 2010). See figure 9.

A PDF (probability density function) graph is similar to a histogram: it is of a random variable and is a function which describes the density of probability at each point in the sample space (De Kant, 2011)

3.1.2. ASQ background

This valuable diagnostic tool allows the clinician to perform comprehensive ultrasound diagnostic in specific regions of the body, such as the abdomen (Toshiba Medical Systems, Europe).

Toshiba Medical Systems Europe introduced the new ASQ technology in Zoetermeer in the Netherlands in October 2009. It is a newly developed advanced clinical application that can analyse echo signal data at more than 100 times higher resolution than normal greyscale or colour flow images using a special hardware extension (Toshiba Ultrasound raw data interface).

Naohisa Kamiyama, the designer of ASQ, explained that most ultrasound units are unable to accept pure acoustical radio frequency (RF) data: instead, only selected segments of this signal are extracted for display on the typical B-mode image seen on the screen. Moreover, ASQ begins with a signal processor that receives the raw RF signal and extracts data that possibly indicate extremely small structures or fibre strands (Brosky, 2009).

“ASQ analyses the spatial echo patterns in a region of interest selected by the user on a greyscale image acquired during the normal ultrasound examination. ASQ operates in the background with the raw data from this region and extracts a parameter and its complete probability distribution curve, related to the homogeneity or smoothness of the structures reflecting the ultrasound beam sent by the Aplio AG into the body”.

(Explains Dr. Naohisa Kamiyama from Toshiba ultrasound R&D who has designed the ASQ algorithm). The parameter may indicate pathologic changes of tissue, e.g. fibrotic transformation of liver paranchyma. (Medical Physics, 2009 (online))

ASQ is expected to be less subjective and operator dependent than conventional ultrasound imaging. ASQ offer qualitative visual results (parametric imaging) as well as quantitative results (De Kant, 2011).

Since 2009, several studies have tested the efficiency of ASQ on liver diseases. Toyoda and Kumada 2009 conducted a study in late 2009 that reports the relationship between data analysis by ASQ and liver histology results for 148 cases. They

confirmed that ASQ results closely match histological classification of patients (Toyoda et al., 2009).

In early 2013, a study was performed in China to assess the diagnostic value of ASQ technology in both homogeneous fatty liver and chronic hepatitis B infected liver. A total of 205 patients were included in this study. The researchers in this study concluded that ASQ quantitative parameters can reflect the severity of homogenous fatty liver to a certain extent (Wang et al., 2013). Later in the same year, a similar study was done in Italy but on hepatic fibrosis. ASQ diagnostic accuracy was compared against liver biopsy and it was concluded that ASQ is a promising new technology which offers encouraging results in the diagnosis of both liver cirrhosis and fibrosis. Nevertheless, to date, it has not reached sufficient diagnostic performance to replace current methods (Ricci et al., 2013).

A similar study to test the efficacy of ASQ was carried out on Japanese patients with hepatic stenosis, with a relatively small sample of 42 patients. The results suggest that ASQ is a useful technology for evaluating tissue characteristics of hepatic stenosis and its follow-up (Onodera, 2013).

However, no research has been done on the reliability and reproducibility of this new technique. Therefore, this chapter will test ASQ for repeatability and reproducibility and determine factors that affect ASQ output.

3.2. Repeatability and reproducibility (Phantom studies)

The objective of this section is to determine the variability and reliability of image ASQ caused by random variation during image acquisition, which will affect the reliability of the ASQ technique. This will include three subsections, which are to determine:

- I. The repeatability (consistency) of the image produced due to variation caused by the ASQ software.
- II. The repeatability (consistency) of the image produced due to the random variation caused by the operator.
- III. The reproducibility (agreement) of the images acquired by two operators under identical conditions: by how much are the two likely to differ?

Intra- and inter-operator repeatability and reproducibility are important variables that permit investigation and examination. The reliability of the imaging technique may affect the accuracy of the diagnosis: therefore, it is essential to understand the reproducibility and repeatability of the ultrasound image (Li et al., 2004). In addition, it is necessary to confirm that a single operator can obtain the same results when repeated measurements are made using the same method under identical conditions, as asserted by (Bailey et al., 2007). Since ultrasound is an operator-dependent modality, the degree of variation caused by the operator (intra-operator) must also be recognised to ensure that the application of the ASQ technique in ultrasound is reliable.

The B-mode ultrasound image is exposed to many confounders, which include instrumentation, scanning and reading protocols and operator variables. Therefore, the validity and reliability of imaging assessments are very important aspects for use in routine clinical practice and research.

In this study, repeatability refers to the reliability of the instrument (ASQ software) and the operator in producing a consistent image (in terms of its texture) when repeated within a short period of time under identical conditions, while reproducibility refers to the agreement between two operators performing a similar scan.

3.2.1. Materials, methodology and statistical analysis

The aim of this section is to describe the key topics related to this study. It is divided into two main sections. The first section describes the materials used in this study with an overview of the equipment and the statistical tests used. This includes Ultrasound principles and Transvaginal transducer. The second section discusses the methods used and the procedures followed to accomplish this study, and then briefly explains both the Coefficient of Variation (CoV) and the Bland-Altman plot, which are the statistical methods used in this section.

3.1.3.1. Materials

3.1.3.1.1. Ultrasound scanner:

Ultrasound is an effective and safe investigation tool that can answer a number of clinical questions without the use of ionising radiation (Guy, 2008). It has the advantage of being non-invasive, safe, readily available and inexpensive when compared to other imaging modalities (Lee et al., 2005). Moreover, Barnett (2000)

disclosed that in general, simple B-mode imaging does not produce harmful temperature increases in tissue.

Ultrasound scanner machines were first used in the medical field in the late 1960s, and it is becoming increasingly important as a diagnostic imaging device. Toshiba Aplio (Toshiba Medical System, Europe) was used to scan the participants and acquire the images for this study (see Figure 12). The specifications of this scanner are given in Appendix I.



Figure 12: Ultrasound machine by Toshiba.

3.1.3.1.2. Ultrasound Transducer:

An ultrasound scanner works by sending high frequency sound pulses into the patient's body. The sound waves travel through the patient's body, passing through different types of tissue. The speed of sound varies with different tissue types: for

instance, the average speed of sound through human tissues is 1540 m/s; the speed of sound through fat is 1459 m/s and it passes through bone at 4080 m/s. Whenever sound encounters two adjacent tissue types with different acoustic properties, a proportion of the sound energy is reflected. These boundaries between different tissue types are called *acoustic interfaces*. The physical property of tissue that describes how much resistance an ultrasound beam encounters as it passes through a tissue is called Acoustic impedance. It depends on: the density of the tissue (d , in kg/m^3) and the speed of the sound wave (c , in m/s), and they are related by:

$$\mathbf{Z} = \mathbf{d} \times \mathbf{c} \quad \text{Equation 8: Acoustic impedance}$$

So, if the density of a tissue increases, impedance increases and if the speed of sound increases, then impedance also increases. The ability of an ultrasound wave to transfer from one tissue type to another depends on the difference in impedance of the two tissues. If the difference is large, then the sound is reflected (Rumack, 2005)

The scanner calculates the distance from the probe to the acoustic interfaces based on the time taken for echoes to return to the probe. The distances and intensities are then displayed on a screen to form a two-dimensional (2D) image (see Figure 13).

ULTRASOUND SYSTEM BLOCK DIAGRAM

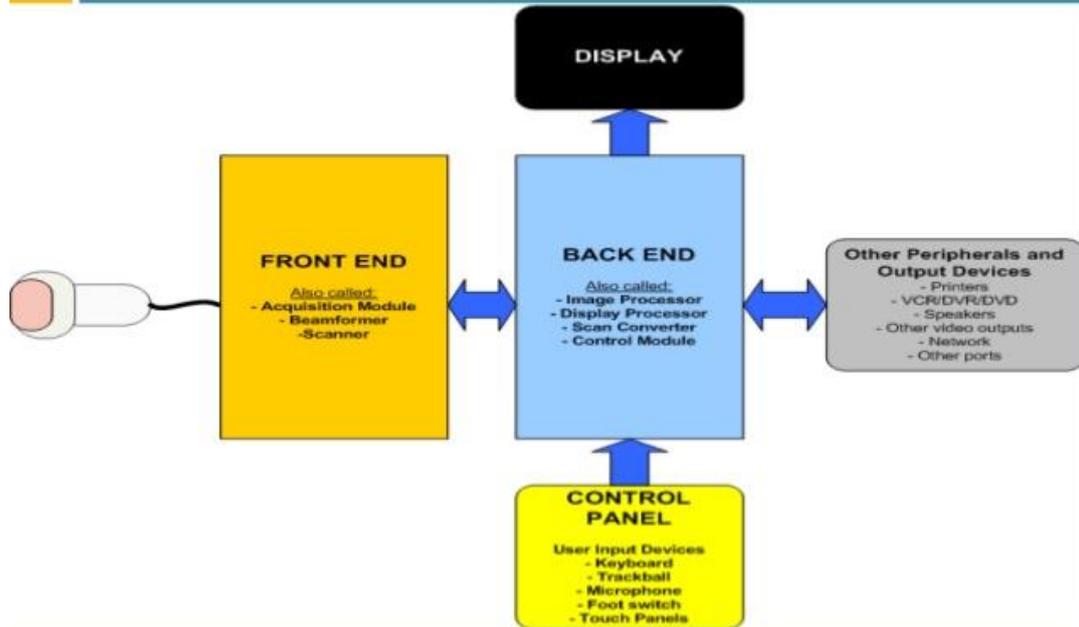


Figure 13: how an Ultrasound scanner works

An ultrasound image is composed of array of pixels of different density. Each pixel represents a discrete intensity that results from the reflection of the ultrasound beam. An image is composed of thousands of pixels. It is represented by one of 256 (0 to 255) shades of grey ranging from black (represented as '0') to white (represented as '255') – (Sanders, 1998).

The transducer probe is any device that converts one form of energy to another. In the case of ultrasound, the transducer converts electrical energy to mechanical energy and vice versa. In other words, it is the part of the machine that produces the sound waves and receives the echoes (Rumack, 2005, Sanders, 1998). It consists of one or more crystals in a plastic housing. These crystals are called piezoelectric (PZ)

crystals. Transducers used in medical ultrasound employ the piezoelectric effect to generate sound waves and detect echo signals. The piezoelectric effect was discovered in the 1880s. Langevin found that when a force is applied perpendicular to the faces of quartz crystal, an electrical charge results (Rumack, 2005). This charge can be detected and amplified, producing a useful electrical signal. Conversely, if an electrical signal is applied to the crystal, the crystal vibrates, sending a sound wave into the medium: hence the dual action of the piezoelectric transducer as a detector and transmitter of acoustic signals (Zagzebski, 1996, Rumack, 2011). This signal is then processed to form the ultrasound image and displayed on a screen. The same piezoelectric crystals are used for sending and receiving the ultrasound pulses: therefore, they have to operate in a switched or pulsed mode. This means that they emit a quick sound pulse, rest and then listen for the echo. This switching between transmitting and receiving modes happens many thousands of times a second (Sanders, 1998). The most commonly used crystals are lead zirconate titanate. The matching layers (plastic nose) lie in front of the transducer element and provide an acoustic impedance difference between the transducer element and the skin. A damping material such as rubber is attached to the back of the transducer element to decrease secondary reverberations of the crystal with the returning signals. Decreasing the ring time results in an increase in depth resolution – see Figure 14.

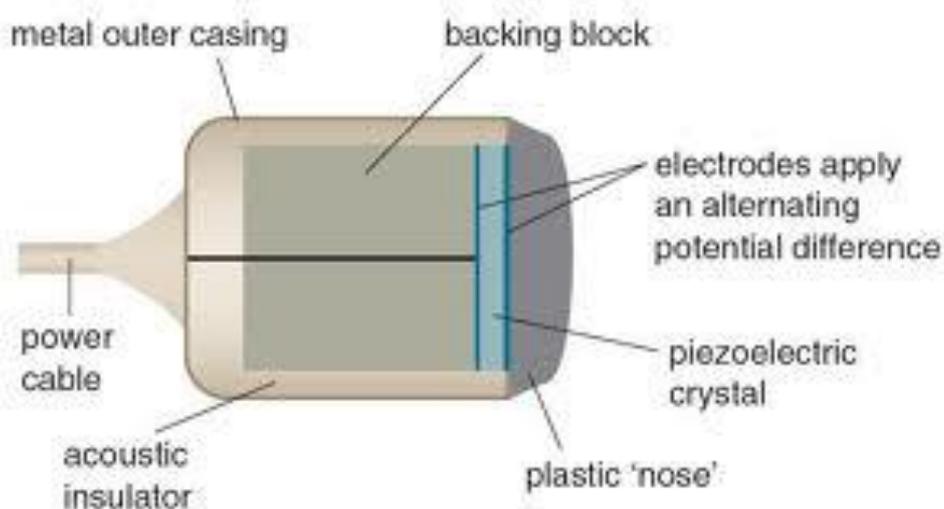


Figure 14: Ultrasound Probe components

Ultrasound probes are designed so that they can be positioned as close to the subject tissue as possible so that resolution can be kept as high as possible. For this reason, specially designed probes are available that can be inserted through different cavities of the body, such as the oesophagus for examination of the stomach, the vagina for examination of the uterus and the ovaries and the rectum for examination of the prostate gland.

The Transvaginal transducer is widely used for scanning the female reproductive system in general and for ovarian cancer diagnosis in specific. It has several advantages over the Transabdominal transducer, as summarised in Table 7. The Royal College of Obstetricians and Gynaecologists suggested in its guidelines that ovarian cysts should normally be assessed using Transvaginal ultrasound: the reason for this is that it provides superior quality (more detail) than the Transabdominal method (RCOG, 2003).

In the present study, a 6.0 MHz Toshiba Transvaginal transducer model PVT-661VT was used to scan the participants (see Figure 15)



Figure 15: Transvaginal Ultrasound probe

Advantages of Transvaginal Sonography

- Use of higher-frequency transducer with better resolution.
- Examination of patients who are unable to fill their bladder.
- Examination of obese patients.
- Evaluation of retroverted uterus.
- Better distinction between adnexal masses and bowel loops.
- Better characterisation of the internal characteristics of a pelvic mass.
- Better detail of a pelvic lesion.
- Better detail of the endometrium.

Table 7: Advantages of transvaginal transducer (Rumack, 2011)

According to (Lagalla and Midiri, 1998), the quality of an ultrasound image is associated with its ability to represent the real anatomy of the structures being examined as closely as possible. This capability depends on the axial spatial resolution, the lateral spatial resolution and the contrast resolution.

The axial spatial resolution refers to the capability of the ultrasound system to represent separately objects which are arranged along the axis of the beam. This parameter is highly dependent on the transducer frequency: the higher the transducer frequency, the better the axial resolution. However, this causes a decrease in the depth of beam penetration.

The lateral spatial resolution refers to the system's ability to discriminate between two objects situated at the same depth perpendicular to the ultrasound beam. This parameter is dependent on the dimension of the US beam and it can be improved by using a narrow beam (Lagalla and Midiri, 1998).

The contrast resolution refers to the US system's ability to distinguish differences between neighbouring tissue regions. This parameter is dependent on the spatial resolution and the image noise.

Similar to other types of endo-cavity transducer, the advantage of the Transvaginal transducer is that it can be placed in close proximity to the organ under investigation (please refer to Table 7). Hence, there is less attenuation from overriding tissue, and a higher frequency can be used to increase the axial resolution. Furthermore, image distortion and artefacts due to tissue heterogeneity or a strongly reflecting interface between the transducer and the organ are also reduced (Whittingham and Martin, 2010).

3.1.3.2. Methods

An original guide document by the National Physical Laboratory (2005) was used for the experimental setup for the repeatability test. It was titled: “Measurement Good Practice Guide No. 52” (Laboratory, 2005).

“The repeatability of the measurement can be quantified in either of 2 ways:

-
- I. *Set the equipment to perform a number of consecutive tests (e.g. 10) using identical conditions and without removing the specimen between measurements. In this case, the only variables are those relating to the performance of the measurement system and associated statistical (random) effects in data capturing and analysis by the software. For simplicity, this will be referred to as ‘instrument-only repeatability’.*

 - II. *As above, but with the specimen removed completely in between measurements. In this case, effects due to variability in the test set-up are introduced. This will include resetting the device and the precision of repositioning the sample. This will be referred as the ‘instrument-operator repeatability’“*
-

3.1.3.2.1. Intra-operator Repeatability

The objective of this section is to assess the reliability (consistency) of the images produced in ASQ due to the random variation caused by the operator.

The operator acquired a series of ten images using the Toshiba Aplio scanner using three different transducers; the 6.0 MHz Transvaginal transducer (model PVT-661VT), the linear transducer (model PLT-7045BT) with 7.5 MHz and the curve transducer (model PVT-375BT) with 3.5 MHz. The transducer was removed from the test object in between scans, where 10 images are acquired for each transducer. Please refer to Figures 17, 18 and 19 for images of the different types of transducers.

An RMI phantom model 403 GS LE was used as a test object (see Figure 20). The machine setting was kept constant across image acquisition processes; it was set to the manufacturer's pre-defined setting: *Endo-Vaginal-Gynae*. The ASQ button was enabled before taking the images. They were then saved as raw data images on the machine. The images then were transferred to the PC in DICOM files for further analysis. A large ROI (see Figure 16) was drawn in the centre of the image using PC-ASQR version 1.11R001 software.



Figure 16: ASQ window with large ROI

A coefficient of variation (CoV) was calculated for each transducer. The ultrasound repeatability CoV reflects the variability of the images due to random variation in the scans caused by the operator. The lower the COV, the smaller the variation between repeats, and therefore the higher the repeatability (Bailey et al., 2007).

To assess the reliability of the ASQ software, ten consecutive ROI were drawn in the same image in the ASQ window after images were transferred as a DICOM file to a PC. The Mean (μ) and Standard Deviation (SD) were calculated as well as the CoV for each transducer image.



Figure 17: Linear Probe

Figure 18: Convex probe



Figure 19: Transvaginal Probe (TV)



Figure 20: Images of RMI phantom model 403 GS LE

3.1.3.2.2. Inter-operator Reproducibility

To assess the reproducibility, all scanning procedures were performed by two operators. The first operator performed the scan, and after he finished, the second operator performed his scan. Between scans, the transducer was repositioned. This procedure was repeated for the three different transducers. All machine variables were kept constant, such as transducer frequency, focal position and gain setting, as well as the depth and size of the ROI when using the ASQ software to draw the ROI.

3.1.3.2.3. Coefficient of Variation (CoV)

The standard deviation (SD) of a set of measurements is often used to describe the variability by expressing the SD as a percentage of the mean. The resulting measure is called the Coefficient of Variation (Armitage et al., 2009, lang and Secic, 2006a). See equation 8.

$$\text{CoV} = (\text{SD}/\text{mean}) \times 100\%$$

Equation 9: Coefficient of Variation

The Coefficient of Variation (Cov) is used to measure the consistency of the data (the uniformity in the data from the mean) (Bruton et al., 2000). It has the advantage of being independent of the units of observation (Altman, 1995). Furthermore, the CoV is widely used to measure random error (Dudely, 2010). In this study, CoV was used to determine the variation (consistency) in the ASQ software repeatability, intra-operator repeatability and influence of the ROI size variance on the mean.

3.1.3.2.4. Bland-Altman plot

The Bland-Altman plot is used for assessing the agreement between two methods of clinical measurement. It is a scatter plot with the difference of the two measurements for each sample on the vertical axis and the average of the two measurements on the horizontal axis. Three horizontal reference lines are superimposed on the scatter plot: one of them at the average difference between means (x-axis values), along with two

lines to mark the upper and lower control limits of plus and minus 1.96σ respectively, where σ is the SD of the measurements mean differences. The graphical approach that plots differences against means is the most informative approach (Bassani et al., 2007). The presentation of the 95% limits of agreement is for visual judgement of how well two methods of measurement agree: the smaller the range between these two limits, the better the agreement is (Myles and Cui, 2007).

In this study, the Bland-Altman plot was used to show the amount of disagreement between two different operators and to represent reproducibility visually.

3.1.3.3. Results

3.1.3.3.1. *Intra-operator repeatability*

Ten consecutive images were acquired from the test object. The CoVs for the three different transducers ranged from 1.4% to 2.2%, with the linear probe showing the smallest variation (1.4%) and the curved probe showing the largest variation (2.2%). Table 8 summarises the mean, standard deviation (SD) and the CoV for the three different transducers used.

Probe Type	Mean	SD	CoV (%)
Linear	102.4	1.4	1.37%
Curve	96.1	2.1	2.2%
TV	101.7	1.6	1.6%

Table 8: Intra-operator repeatability (different images).

3.1.3.3.2. *Repeatability of ASQ software*

Ten consecutive ROI were drawn on the same image in the ASQ window. The CoVs for the three different transducers range from 0% to 0.32% with the linear probe showing the largest variation (0.32%), while both the TV and the curved probe showed no variation (0%). Table 9 summarises the mean, standard deviation (SD) and the CoV for the three different transducers used.

Probe Type	Mean	SD	CoV (%)
Linear	100.1	.32	0.32%
Curve	93	0	0 %
TV	102	0	0%

Table 9: ASQ software repeatability (same image)

3.1.3.3.3. *Inter-operator reproducibility*

The inter-observer mean differences SD between paired images for the three different transducer types tested are presented in Table 10. P-values were calculated using the Mann-Whitney U test as the data were non-normally distributed. For all three probes, there were no significance differences between operator 1 and operator 2, with p-values of .436 for the linear probe and .579 for both the curved and the TV probes.

Bland-Altman plots were drawn for the three probes and tested for graphic evaluation of the agreement between ASQ feature derived from the images obtained by the two operators. Figures 21, 22 and 23 are examples of Bland-Altman plots showing the distribution of the mean values in the three probes. All probes show different patterns: i.e. there is random variation in the difference between the two operators. The difference in the paired derived ASQ means between the two operators do not exhibit any systematic pattern, as the mean value increases for all three probe types. In other words, the difference between the two operators does not increase (or decrease) systematically as the mean value increases.

The mean CoV for inter-operator repeatability are 1.03, 1.2 and 2.1% for the linear, curved and TV probe respectively. All three probes show very low variation, with values below 2.5%. The mean CoVs for all three probes are presented in Table 11.

	Linear Probe	Curve Probe	TV Probe
Mean Diff	-.60	-.80	-1.0
Std. Dev.	2.27	4.44	4.87

Table 10: Inter-operator mean difference and standard deviation

Mean Coefficient of Variation (CoV %)		
Linear Probe	Curve probe	TV probe
1.03	2.1	1.2

Table 11: Coefficient of variation for inter-operator reproducibility.

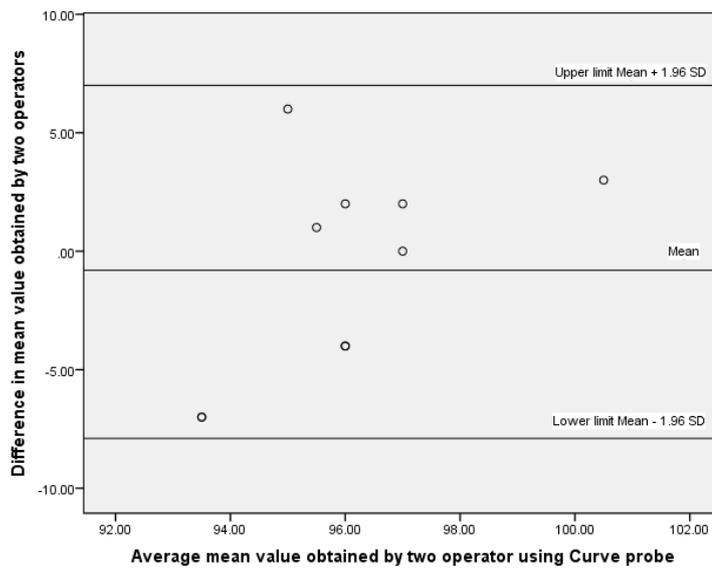


Figure 21: Example of Bland-Altman plot showing the random distribution of the difference between the two operators (curve probe)

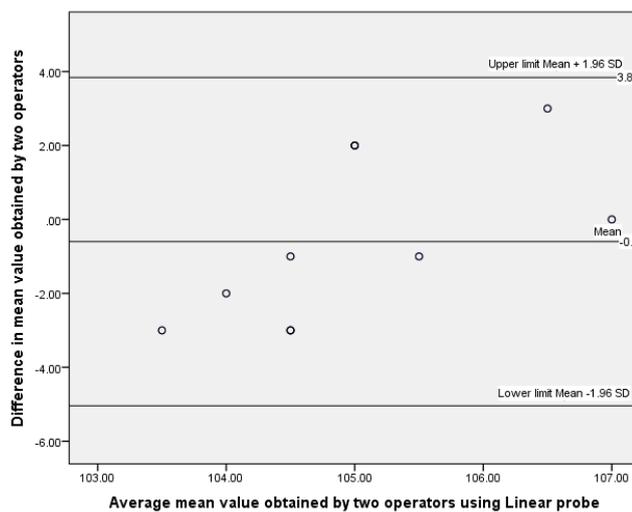


Figure 22: Example of Bland-Altman plot showing the random distribution of the difference between the two operators (linear probe).

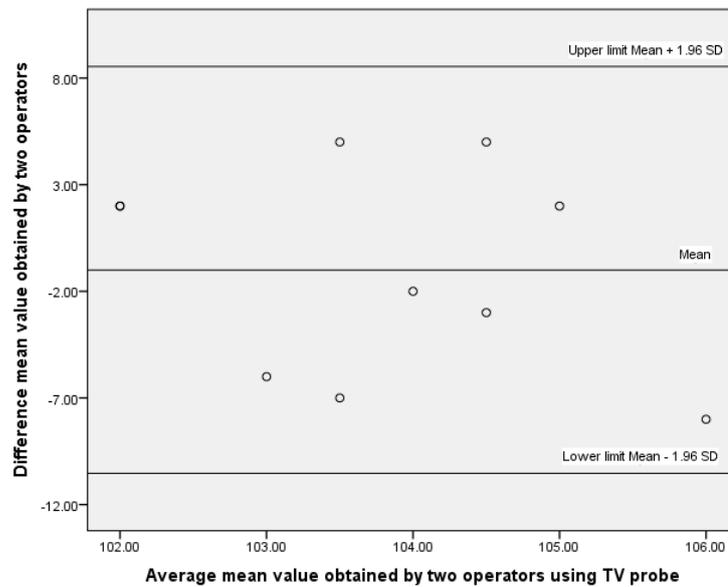


Figure 23: Example of Bland-Altman plot showing the random distribution of the difference between the two operators (TV probe).

3.1.3.4. Discussion

Repeated measurements are most likely to vary (Swinscow, 1997). The degree of the variation will determine the reliability of the measurements acquired. Bruton et al. (2000) specified that reliability refers to the consistency or repeatability of such measurements, and is usually performed to assess the performance of the instrument and the operator. De Vet et al. (2006), pointed out that reliability is an essential requirement for measuring outcomes in medical disciplines, such as the assessment of radiographs.

In this chapter, a set of experiments were carried out to evaluate the reliability of B-mode images with the ASQ feature enabled, particularly in the random variation inherited from the ASQ software and from the operator. Since the reliability of the measurement (image acquisition) determines the sensitivity and the specificity when

used for classifications or predictions, establishing the degree of these variations is essential. The scatter (variation) in the extracted ASQ feature was evaluated as a measure to determine reliability of the image texture. According to Fry (2002), the scatter in the data is a useful measure for quantifying the effect each parameter has on the accuracy of the measurement. The variation caused by image acquisition procedure can be further divided into ASQ software variation and operator variation.

Generally, ASQ features in all three transducers evaluated demonstrated excellent repeatability for ASQ software, with all transducers showing less than 0.4% variance from the mean. The results indicate that ASQ software is able to produce reliable ASQ output measures.

In this section, we also assessed the intra- and inter-operator reliability in image acquisition. It was observed that for all types of transducers, their intra-operator CoVs are 1.37% - 2.2%, which are higher than the ASQ software CoVs. This can be explained by the fact that the variations in the scanning process are a combination of both the ultrasound system and the operator. However, CoV of 2.2% or less is considered very low and indicates that this is a reliable technique.

Regarding inter-operator reproducibility, the Bland-Altman plots clearly illustrate that the differences between the two operators did not appear to increase as the mean value of the ASQ increases using the linear probe and are due to random variation. The same pattern was observed for the two other probes: the curved and TV probes. In other words, there is no obvious relationship noted between the differences of the two

operators and their means for all probes evaluated. This indicates that the increases in the derived ASQ mean values do not affect the difference between the two operators. According to Bland and Altman (2010), if the differences are proportional to the mean, a logarithmic transformation of the data has to be performed before analysis is carried out.

The CoV calculated to measure the consistency of the image produced by two different operators shows a very small variation (excellent agreement).

From the results above, it can be concluded that the ultrasound system with ASQ features is able to produce consistent images. Since the results show that scanning using a phantom is able to produce consistent images, it is worth continuing to an assessment of other factors that might influence the reliability of ASQ output measurements. This is addressed in the next section.

3.3. (B) Influence of ROI size, ROI depth, Focus, Gain setting and transducer frequency on ASQ

The objective of this section is to assess factors that potentially affect the ASQ curve, namely the ROI size, ROI depth, focus, gain setting, and transducer frequency. Subsection 1 reviews the influence of those five factors on the ASQ curve. Subsection 2 contains the methodology and statistical analysis, while subsection 3 presents the results, which are discussed in subsection 4. Finally, subsection 5 summarises this chapter.

3.3.1. Background

It was reported that besides the variation in the imaged tissue itself, the output measurement may also vary due to other factors. Since there were no available articles or publications on ASQ image properties, studies of texture analysis were considered as a reference for background information for ASQ. This is due to the similar nature of ASQ and texture analysis in the fact that they are both statistically analysed by computer software.

According to Castellano et al. (2004), the effect of external factors on some texture parameters must be taken into consideration before using texture analysis techniques. If this is not addressed, then the value obtained may not reflect the actual texture of the tissue, which may in turn influence the performance of the texture analysis technique when applied for classification, as pointed out by Chan and McCarty (1990).

In previous work on a tissue-equivalent phantom, it was reported that some texture features exhibit a dependency on the size of the area from which the features were extracted (He et al., 2004).

In the same way, the influence of the ROI size has been reported in clinical image studies. For example, a study was done to discriminate prostate tissue from the ultrasound images. It was found that results from using smaller ROI were not as good as results using larger ROI. This demonstrates that the size of ROI will affect the performance of the extracted feature (Basset et al., 1993). Furthermore, He et al. (2004) argued that determining the appropriate ROI size is crucial due to the variation that might be in the pathological area and that the tumour size is generally larger than the size of the normal area.

Beside the ROI size, another factor that needs to be considered when acquiring the ASQ images is the depth of the ROI. In 1988, Morris conducted a study which revealed that some features of the texture analysis are dependent on the ROI depth (Morris, 1988). To overcome this limitation, some studies used a fixed ROI depth. However, this approach restricts the flexibility of the technique because real-life lesions may appear at various depths.

It has been reported by several studies that even when imaging the same tissue, the value of the extracted features may vary under different scanner settings (Collewet et al., 2004, Chan and McCarty, 1990). Gain and focus position are examples of such factors. Therefore, in this section, the five main factors that may influence the

performance of the ASQ will be tested in a set of experiments to acknowledge the factors that have an effect on ASQ and to use them with caution in tissue studies.

3.3.2. Methodology and statistical analysis

3.3.2.1. Influence of ROI size

The objective of this section is to evaluate the influence of the ROI size on the ASQ feature, particularly on the mean of the output curve.

Ten images were acquired using a Toshiba Aplio 500 scanner with three different transducers: a 6.0 MHz Transvaginal transducer (model PVT-661VT), a curve transducer (model PVT-375BT) at 3.5 MHz and a linear transducer (model PLT-7045BT) at 7.5 MHz (as seen in Figures 17, 18 and 19 respectively). An RMI model 403 GS LE phantom was used as a test object (please refer to Figure 20).

The machine settings were kept constant by using the manufacturer's pre-defined setting: *Endo-Vaginal-Gynae*. The ASQ button was enabled in the machine before taking the images. The set of ten images were saved as raw data on the ultrasound machine hardware and then transferred to the PC in a DICOM file for further analysis. Figure 24 illustrates the experimental setup.

ASQ software, PC-ASQR Version 1.11R001, was used to draw the different ROI sizes on the images and acquire the output curve. The same process was repeated for the three different transducers. The mean of the output curve was documented from two different ROI sizes: large ROI and small ROI. The significance of the difference

was calculated in each transducer using the Mann-Whitney U test p-value to test for significance because of the non-parametric data found when tested for normality.

3.3.2.2. Influence of ROI Depth

The objective of this section is to evaluate the influence of the ROI depth on the ASQ; whether as the depth of the ROI increases, the value of the mean of the output curve tends to increase/decrease as well. The procedure carried out for the image acquisition was similar to the ROI size assessment (see previous paragraph). The focus was kept constant in all images, while the ROI were drawn at two different depths (2cm and 4 cm) in each image. The p-value were calculated to determine the relationship between the ROI depth and the ASQ mean of the output curve, i.e. whether there is a significant difference when changing the ROI depth regarding the focus.

3.3.2.3. Influence of Gain setting

The objective of this section is to evaluate the influence of the scanner gain setting on the ASQ output curve, i.e. whether there is a significant difference of the mean when changing the gain setting.

Ten images were acquired using a Toshiba Aplio 500 scanner with three different transducers; a 6.0 MHz Transvaginal transducer (model PVT-661VT), a 3.5 MHz curved transducer (model PVT-375BT) with and a 7.5 MHz linear transducer (model PLT-7045BT). An RMI model 403 GS LE phantom was used as a test object. The machine settings were kept constant by using the manufacturer's pre-defined setting:

ASQ setting (except for the gain). The ASQ button was enabled in the machine before taking the images. The images were acquired at two different gain settings: 85 % (low) and 100 % (high). The set of ten images were saved as raw data on the ultrasound machine hardware and then transferred to the PC in a DICOM file for further analysis. Please refer to Figure 24 for the experimental setup and the process flow illustration. A relatively large ROI was drawn in the centre of the image for the output curve using the ASQ software version 1.11R001.

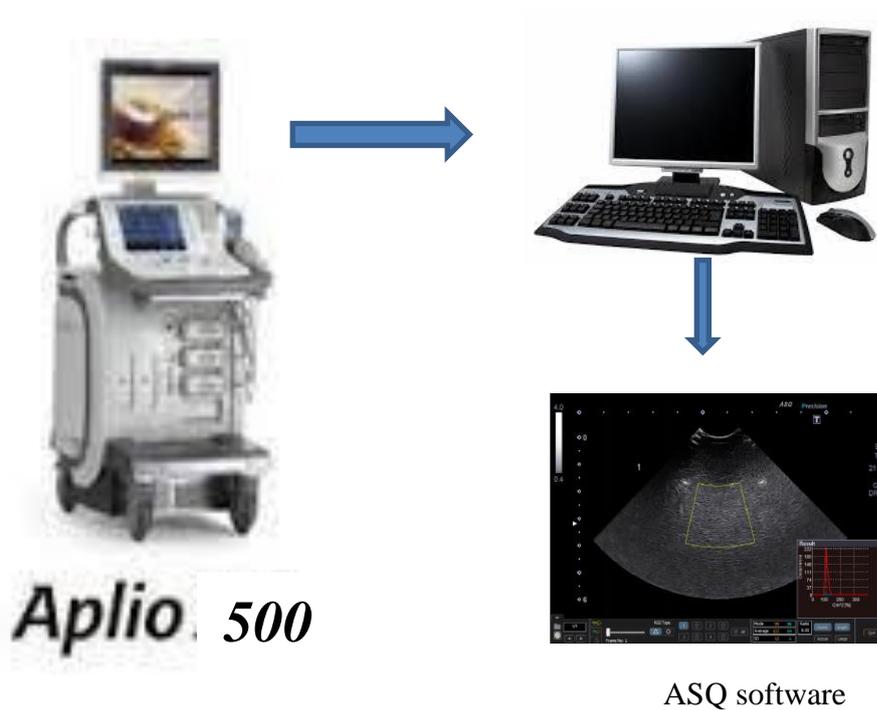


Figure 24: Experiment process flow illustration.

3.3.2.4. Influence of focus position

The objective of this section is to evaluate the influence of the focal position in the ultrasound image on the ASQ output curve, i.e. whether there is a significant difference of the mean when changing the focal position.

The same procedure was followed with the same ultrasound machine and type of transducers used as in the gain setting experiment (see previous section). The only part of the procedure that was different is that the gain was kept constant along with the other factors and ten images were acquired with two different focal positions, one at 2 cm and the other at 4 cm.

3.3.2.5. Influence of transducer frequency

The objective of this section is to have an understanding of the influence of transducer frequency on the ASQ output curve.

A similar procedure was followed using the same ultrasound machine and type of transducers as in the gain setting experiment (see gain section). All variables were kept constant, i.e. the gain setting, focal position, depth of ROI and size of ROI. The only difference here is the frequency of the transducer. For the linear transducer, three different frequencies were tested: 8.4, 7.2 and 6.2 MHz; for the curved transducer, four different frequencies were tested: 6, 5, 4 and 3 MHz. For the TV transducer, three frequencies were tested: 8, 7.2 and 6 MHz. For each frequency, ten images were acquired and saved as raw data. These images were then transferred to the PC in a

DICOM file for further analysis. A relatively large ROI was drawn in the centre of the image for output curve using the ASQ software version 1.11R001.

3.3.3. Results

3.3.3.1. Sample characteristics

A Shapiro-Wilk normality test ($p > .05$) and a visual inspection of the histograms, normal Q-Q plots and box plots showed that some of the variables were not normally distributed when calculating the z-value, by dividing the skewness and kurtosis values over the standard error for each of them. Therefore, a non-parametric test (Mann-Whitney U) was used to calculate the p-values for all of the variables to keep the results consistent.

3.3.3.2. Influence of ROI size

The median was calculated for each transducer from the output curve in each image. These results are presented in Tables 12, 13 and 14. These results show that there is a significant difference between the large ROI and the small ROI in the linear transducer with p-value $< .001$, while in the curved and TV transducers, the difference is not significant, with p-values of 0.105 and 0.912 respectively (see Figures 25 and 26).

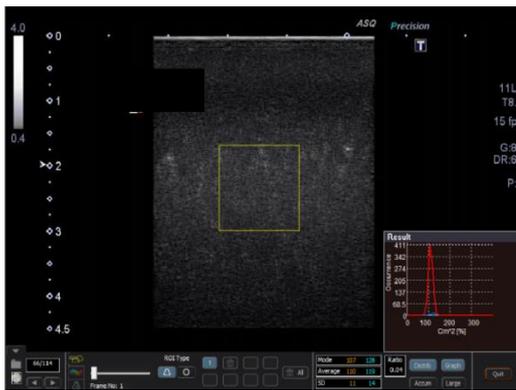


Figure 25: ASQ window with large ROI

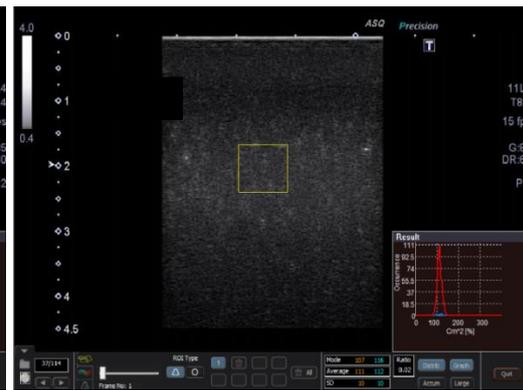


Figure 26: ASQ window with small ROI

	Median	p-value	Significance?
Large ROI	110	<.001	Yes
Small ROI	112		

Table 12: Summary of the influence of the ROI size on the mean of output curve using Linear Probe

	Median	p-value	Significance?
Large ROI	98.50	0.105	No
Small ROI	97.50		

Table 13: Summary of the influence of the ROI size on the mean of output curve using Curve Probe.

	Median	p-value	Significance?
Large ROI	103	0.912	No
Small ROI	103		

Table 14: Summary of the influence of the ROI size on the mean of output curve using TV Probe.

3.3.3.3. Influence of the ROI depth

The median was calculated for each transducer due to the non-normal distribution of some of the variable data from the output curve in each image. These results are

presented in Tables 15, 16 and 17. Furthermore, it shows that there is no significant difference for either the curved or the TV probe, with p-values of .769 and .631 respectively, when placing the ROI at depths of 2cm or 4cm. On the other hand, the linear probe showed a significant difference in the median when drawing the ROI at 2 cm and 4 cm with p-value <.001 using the Mann-Whitney U test (see Figures 27 and 28).

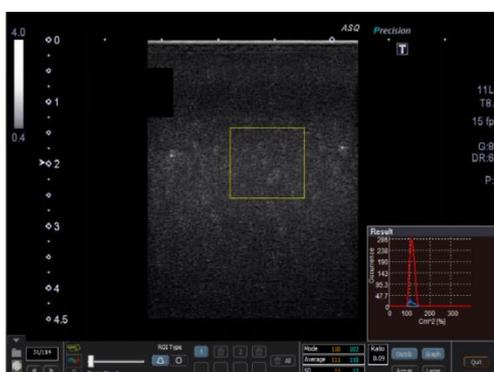


Figure 27: ASQ window with ROI at 2cm depth

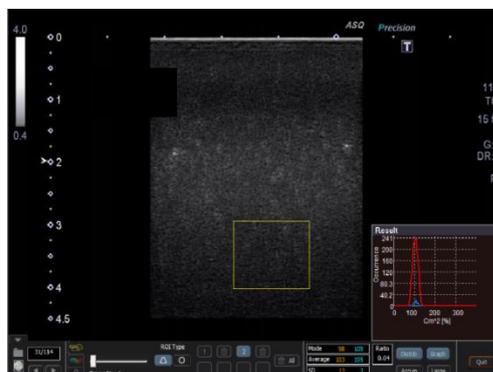


Figure 28: ASQ window with ROI at 4cm depth

	Median	p-value	Significance?
2cm depth	112	<.001	Yes
4cm depth	103		

Table 15: Summary of the influence of the ROI depth on the mean of output curve using linear probe.

	Median	p-value	Significance?
2cm depth	98.5	0.769	No
4cm depth	99		

Table 16: Summary of the influence of the ROI depth on the mean of output curve using curved probe.

	Mean	p-value	Significance?
2cm depth	104	0.631	No
4cm depth	103.5		

Table 17: Summary of the influence of the ROI depth on the mean of output curve using TV probe.

3.3.3.4. Influence of the focus setting

The results of the influence of the focus on the output curve are summarised in Tables 18, 19 and 20. The linear and TV probes showed a significant difference in the output curve mean when changing the focus from 2 cm to 4 cm, with p-values of .003 and <.001 respectively. In contrast, there was no significant difference in the output curve mean when changing the focus from 2cm to 4cm in the curved transducer, with a p-value of .579.

Figures 29 and 30 illustrate how the images were taken with different focal positions prior to saving and transferring them to PC in a DICOM file.

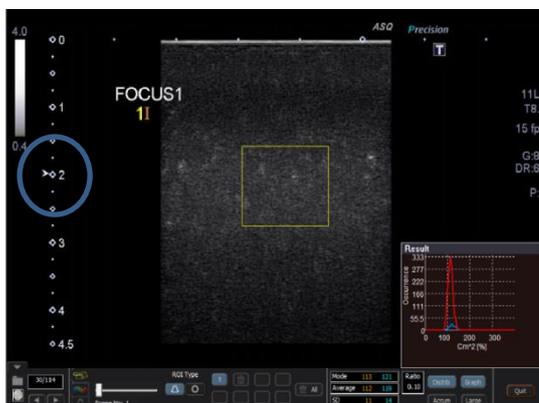


Figure 29: ASQ window with focus at 2 cm.

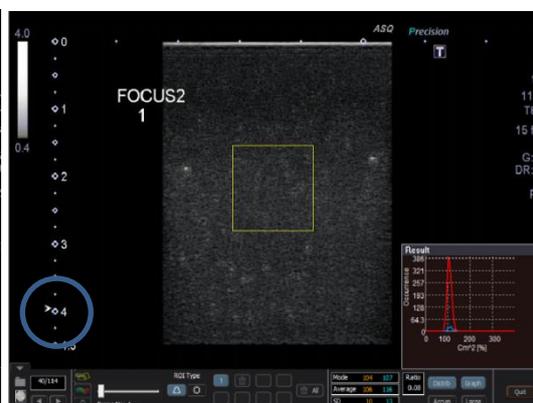


Figure 30: ASQ window with focus at 4 cm.

	Median	p-value	Significance?
Focus at 2cm	112	0.003	Yes
Focus at 4cm	109		

Table 18: Summary of the influence of the Focus setting on the mean of output curve using linear probe

	Median	p-value	Significance?
Focus at 2cm	98.5	0.579	No
Focus at 4cm	100		

Table 19: Summary of the influence of the focus setting on the mean of output curve using curved probe

	Median	p-value	Significance?
Focus at 2cm	103	<.001	Yes
Focus at 4cm	106		

Table 20: Summary of the influence of the focus setting on the mean of output curve using TV probe.

3.3.3.5. Influence of gain setting

Tables 21, 22 and 23 show the results when changing the gain settings on the three different transducers respectively. The median was calculated. In all three transducers, the gain setting did not give any significant difference in the mean of the output curve, with p-values of .393 for the linear probe and .853 for both the curved and TV probes (see Figures 31 and 32).

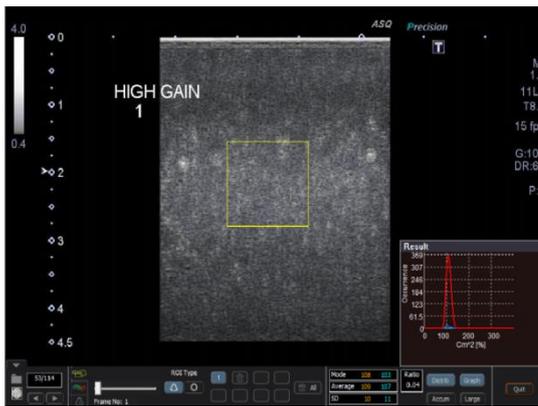


Figure 31: ASQ window with High Gain of 100 %

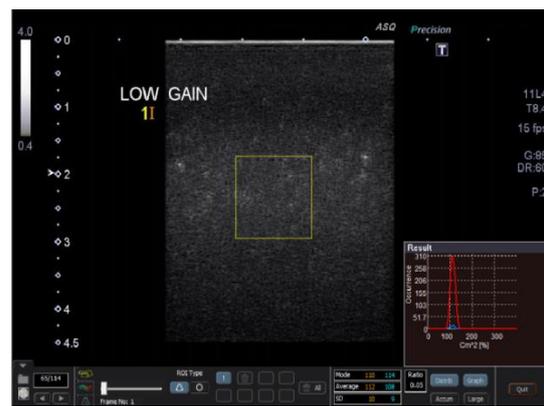


Figure 32: ASQ window with low Gain of 85 %

	Median	p-value	Significance?
High gain	112	0.393	No
Low gain	110		

Table 21: Summary of the influence of the gain setting on the mean of output curve using linear Probe.

	Median	p-value	Significance?
High gain	101	0.853	No
Low gain	101		

Table 22: Summary of the influence of the gain setting on the mean of output curve using curve Probe.

	Median	p-value	Significance?
High gain	102	0.853	No
Low gain	102.5		

Table 23: Summary of the influence of the gain setting on the mean of output curve using TV Probe.

3.3.3.6. Influence of transducer frequency

The influence of frequency on the output curve is summarised in Tables 24, 25 and 26. The mean and median of the ASQ output curve were calculated for all frequencies. For the linear probe, the 8.4 MHz frequency was significantly different when compared to the 7.2 MHz frequency as well as when compared to the 6.2 MHz frequency, with both p-values of $< .001$. However, there was no significant difference seen when comparing the p-values of 7.2 with 6.2 MHz, with a p-value of .739.

For the curved probe, four different frequencies were tested against each other. In all frequencies, no significance difference was seen when compared to each other, with p-values ranging from .165 to .796 using the Mann-Whitney U test.

When testing the TV probe, no significant differences were seen between the frequencies 8 and 7 MHz and between frequencies 7 and 6 MHz, with p-values .280 and .089 respectively. The only significant difference shown was between 8 and 6 MHz, with a p-value of .003.

Frequency (MHz)	Mean	Median	P-value	Significance?
8.4	71.6	72	(8.4 with 7.2) <.001	Yes
7.2	103.6	104	(7.2 with 6.2) .739	No
6.2	103.3	103.5	(8.4 with 6.2) <.001	Yes

Table 24: Influence of frequency in linear transducer.

Frequency (MHz)	Mean	Median	P-value	Significance?
6	96.4	97	(6 with 5) .436	No
5	95.6	95.5	(5 with 4) .739	
4	95	95	(4 with 3) .796	
3	93.8	95	(6 with 3) .353	
			(6 with 4) .165	

Table 25: influence of frequency in curved transducer.

Frequency (MHz)	Mean	Median	P-value	Significance?
8	104.5	105	(8 with 7.2) .280	No
7.2	103.1	104.5	(7.2 with 6) .089	No
6	101.1	101	(8 with 6) .003	Yes

Table 26: influence of frequency in TV transducer.

3.3.4. Discussion

The success of applying the ASQ technique is subject, among other factors, to the reliability of the extracted features. Five factors associated with the reliability were considered in this section: (i) ROI size, (ii) ROI depth, (iii) focus, (iv) gain setting and (v) transducer frequency.

The results for the influence of ROI size on the mean of the output curve show that the size of the ROI does not affect the occurrence curve mean. This can be clearly seen in both the TV transducer and the curved transducer. Yet, in the linear transducer, the variation is inversely proportional to the size of the ROI. In other words, the larger the ROI size, the more reliable the mean of the output curve, with less variation, and the smaller the ROI, the more variation can be seen in the mean. This indicates that the linear transducer is ROI-size dependent. This could be explained by the fact that a smaller number of samples are used in the smaller ROI, which gives more variation. In other words, a small ROI would have an insufficient

numbers of pixels to reliably compute the ASQ mean output. As different lesions may have different sizes, a minimum size that ensures that the variation in ASQ parameters in each ROI is captured needs to be established.

It is important to note that a relatively large ROI size should be used for any particular study using the linear probe to avoid misleading results due to the difference in the ROI size. This transducer, however, was not used in the study of images of ovarian masses.

These results do not apply to the TV or the curved transducer. It has been shown that the ROI size does not affect the mean of the output curve. Moreover, the mean of the output curve will have minimal variation when using different ROI sizes that will not affect its reliability.

ROI depth is another determinant of the mean of the output curve. The ultrasound beam causes a depth dependence of the B-mode image texture, as described by Oosterveld et al. (1991). This can be explained by the fact that the intensity progressively decreases as the beam passes through tissues as a result of scattering, refraction and absorption phenomena (Lagalla and Midiri, 1998).

The attenuation on the ultrasound signal as it advances through the tissue is one of the underlying factors that affect the B-mode image texture. According to Morris (1988), the correction of this variation in the raw image is not straightforward and might not even be possible. To overcome this problem, some studies have confined the ROI

position to a fixed depth to avoid depth dependency. For instance, (Bader et al., 2000) stated in their study that all tumours were located at depths between 2 cm and 3 cm.

Since the position of the lesion may vary, it is crucial to understand the influence of the ROI depth on the mean of the output curve. We have evaluated the effect of ROI depth on the mean of the output curve on three different transducer types. In both the TV and the curved transducers, the depth did not significantly affect the mean of the output curve, while in the linear transducer, a significant difference was seen when using different depths for the ROI.

The third factor that may affect the ASQ output curve is the position of the focus. Generally, the ultrasound beam spreads out, or undergoes divergence, as it moves away from the transducer (International Society of Radiology: (Tole, 2005)); therefore, it is possible to focus the ultrasound beam to cause narrowing of the beam and thus improve (lateral) resolution. The beam can be narrowed at a predetermined distance from the transducer. The point at which the beam is at its narrowest is the focal point or focal zone and is the point of greatest intensity and best lateral resolution (George, 2006).

The fact that the focus position does affect the mean of the output curve in the TV probe led us to decide to keep the focus position at the level of the drawn ROI to acquire the greatest intensity and the best lateral resolution to ensure reliable results.

The fourth factor that may influence the mean of the output curve in the ASQ software is the gain setting. The results of testing the effect of high gain (100 %) and

low gain (85 %) showed that all three probes are independent of the gain setting. This means that using either high or low gain when acquiring the raw data on the ultrasound machine will not affect the mean of the output curve in the ASQ software. This could be explained by the fact that the ASQ software uses the raw data, which is initially not affected by the gain setting.

Transducer frequency results showed that the mean of the output curve is not affected by the frequency chosen to capture the ultrasound images except when choosing 8.4 or 8 MHz in the linear and the TV probe respectively. These higher frequencies seem to affect the mean of the output curve in the ASQ.

Based on the findings from this chapter, three factors do not influence the mean of the output curve in the ASQ software using the TV probe: these are ROI size, ROI depth and the gain setting. Nevertheless, the focal position does have a significant effect on the output curve. Moreover, transducer frequency does affect the mean of the output curve only if a high frequency is chosen (8 MHz); otherwise lower frequencies do not influence the mean of the output curve. These findings permit us to use the ASQ software on ovarian masses with confidence that these factors will not affect the reliability of the reading, while taking caution when positioning the focus on the image before storing the raw data as well as using low transducer frequencies.

3.3.5. Testing ASQ settings

Since ASQ is a new and evolving technique, further exploration was needed to understand the correct procedure to be used. Usually only Toshiba Aplio 500

machines come with an ASQ pre-setting and an ASQ button. However, it was worth testing different Toshiba models to acquire raw images that could be used in the ASQ window and analysed using the ASQ software.

For this purpose, the aim of this test is to compare the Toshiba Aplio 400 to the 500 model as well as to test three different pre-settings on the same machine. Using a linear probe, ten images were acquired in the Aplio 500 and stored as raw data, and then copied as a DICOM file to be analysed using the ASQ software. In the Aplio 400, the machine was set up by a Toshiba engineer to enable raw data store button in the control panel. Then ten images were acquired using a similar technique and stored as a raw file. These images were then transferred as a DICOM file to a personal PC to be analysed using ASQ software.

Then three different pre-settings were tested: these are pre-set off (the raw data) with ASQ on, pre-set on with ASQ on and pre-set on with ASQ off. In each setting, ten images were acquired and stored as raw data, all other variables were kept constant throughout the test. Images then were copied to a flash memory drive as a DICOM file to be analysed using the ASQ software.

3.3.5.1. Results

Two different tests were done to investigate the most reliable setting of the machine to acquire the images before transferring them to be analysed using ASQ. First, the different machine models (Aplio 500 and Aplio 400) were compared when saving images as a raw store without enabling the ASQ button. The p-value was .912,

indicating no significant difference between the two models when storing the images as a raw store without enabling the ASQ button before saving the images. Table 27 summarises the results.

Second, three different pre-settings of the machine were compared: these are pre-set button off with ASQ on, pre-set button on with ASQ on and pre-set button on with ASQ off. Surprisingly, there were no significant differences between pre-set button on and off when the ASQ button was on, with a p-value of .796, while there was a significant difference when the ASQ button was on compared to when the ASQ button was off even when the pre-set button was on, with a p-value <.001. (Please refer to Table 28 for results).

Machine Model	Median	P-value	Significant?
Aplio 500	72	.912	No
Aplio 400	72		

Table 27: The difference between Toshiba model Ultrasound machines

Machine setting	Mean	Median	P-value	Significance?
Pre-set OFF ASQ on	105.1	105	.796	No
Pre-set on ASQ on	104.9	105		
Pre-set on ASQ off	71.3	71	<.001	Yes

Table 28: Summary of the results of ASQ machine settings.

3.3.5.2. Discussion and Conclusion

ASQ software was initially designed to work with Toshiba Aplio 500 machines. However, further testing had to be done on the Toshiba machines to investigate the right setting to acquire images before analysing it using ASQ software. It was possible to acquire images on Aplio 400 after setting a raw store button on the machine. Unfortunately, this gave misleading results when analysed using ASQ software. Because we used a phantom that consisted of uniform tissue, we expected the mean and median of the output curve to be around 100 to represent the uniformity of the tissue, but the mean and median of the output curve were shifted to the 70s.

When testing the machine setting, it was found that the pre-set button on the control panel screen was not significantly affecting the raw stored images when analysed using ASQ software. On the other hand, the ASQ button had a major influence on the image property before storing. This could be explained by the fact that the pre-set

button only enabled the machine to switch on precision; Precision is defined as an adaptive processing algorithm which looks at the returning echo signals and uses image processing techniques to extract the real information from the background noise. This has the effect of reducing the clutter in the image and producing an image with finer detail (Brosky, 2009).

However, the precision button was found to have no effect on the raw store images. Another reason is that the manufacturer of the ASQ has set this button to enable some filters that are not available on older Aplio models.

From these tests, it can be concluded that ASQ software will not work properly and reliably on images taken from machines other than Aplio 500. Moreover, the ASQ button has to be enabled on the machine before storing images in raw form.

3.4. (C) Influence of pre-defined image parameters on ASQ output

In the ASQ software window, there is a bar of multiple pre-defined parameters that can be altered before selecting the ROI. These parameters have a dramatic effect on the output graph shape, average, mode and SD values.

These parameters are the x and y values for the sample chosen, the sample steps, sweep intervals and the total number of samples. The software comes with default settings for these parameters, which are shown in Table 29. However, these default settings were originally created to be used on liver tissue. Therefore, in this section, the default setting will be tested on a phantom used in Ultrasound in an attempt to create a set of parameters that are most useful to be used on images of ovarian masses.

Samples	
x 21	y 19
Sample steps	
x 0	y 2
Sweep intervals	
x 2	y 4
Actual size	
Number of samples 315	

Table 29: The default settings in the ASQ software window

The large ROI is the region that is drawn within the area of interest, while the small ROI is the sample areas, within the large ROI, that are used to determine the cm^2 value, as explained by Toshiba engineer. The size of the small ROIs is determined by the value of X and Y in the toolbar setting.

In other words, when drawing a ROI on an image, it consists of number of small ROI next to each other, and samples x and y represent a pre-set number of scan lines that will be used on both the x axis and the y axis in the small ROI. Within the large ROI, at least 100 small ROIs are needed in order for the statistics to be meaningful. The sample steps represent the number of un-sampled pixels between each sampled pixel in the small ROI. This can be determined in the x and y planes independently. When the SS is 0 in each plane, every pixel in the ROI is used for the analysis. The smaller the sample steps, the greater the number of sampled pixels within the small ROI. The sweep interval determines how close the small ROIs are, as they are being interrogated: in other words, it represents the gap between two small ROI. For our small pathological structures, we need them to be as close as possible, so that we can detect very small areas of speckle variation.

Finally, the number of samples gives the actual number of small ROI that will be calculated when drawing the ROI on an image. It is controlled by the values of x and y. Furthermore, it needs to be adjusted to more than 100, as advised by the Toshiba engineer.

3.4.1. Methodology

The goal of these experiments is to define a set of parameters that can allow us to draw the smallest ROI (e.g. 1 x 1 cm) without affecting the reliability of the output graph and at the same time ensuring reliable values.

Ten images were acquired using a Toshiba Aplio scanner. A linear transducer (model PLT-7045BT) with 7.5 MHz was used for this section to establish new parameter settings to be used later on images of ovarian masses. The reason for this is that the linear transducer has the simplest image and gives a square shaped ROI that is easier to draw and measure.

An RMI model 403 GS LE phantom was used as a test object. The machine settings were kept constant by using the manufacturer's pre-defined setting: *Endo-Vaginal-Gynae*. The ASQ button was enabled in the machine before taking the images. The set of ten images were saved as a raw store on the ultrasound hardware and then transferred to the PC in a DICOM file for further analysis. Figure 33 illustrates the experimental setup.

The same set of ten images were used throughout this section while changing the pre-defined parameters from the left bar shown on the software window (please refer to Figure 33). In the ten images, the ROI were drawn twice: once with the default parameter setting and the other with changing the steps x, y and sweeps x, y to 0, 0 and 1, 1 respectively. The mean and SD were calculated for both sets of data and compared using t-tests to calculate the p-value for the significance of the difference.

As advised by Toshiba engineers, to ensure reliable results, the actual sample size should not fall below 100. Therefore, several different parameters were used to reach the optimum set with at least a sample size of 100. To achieve that, a set of experiments were done (please see the flow chart in Figure 34).

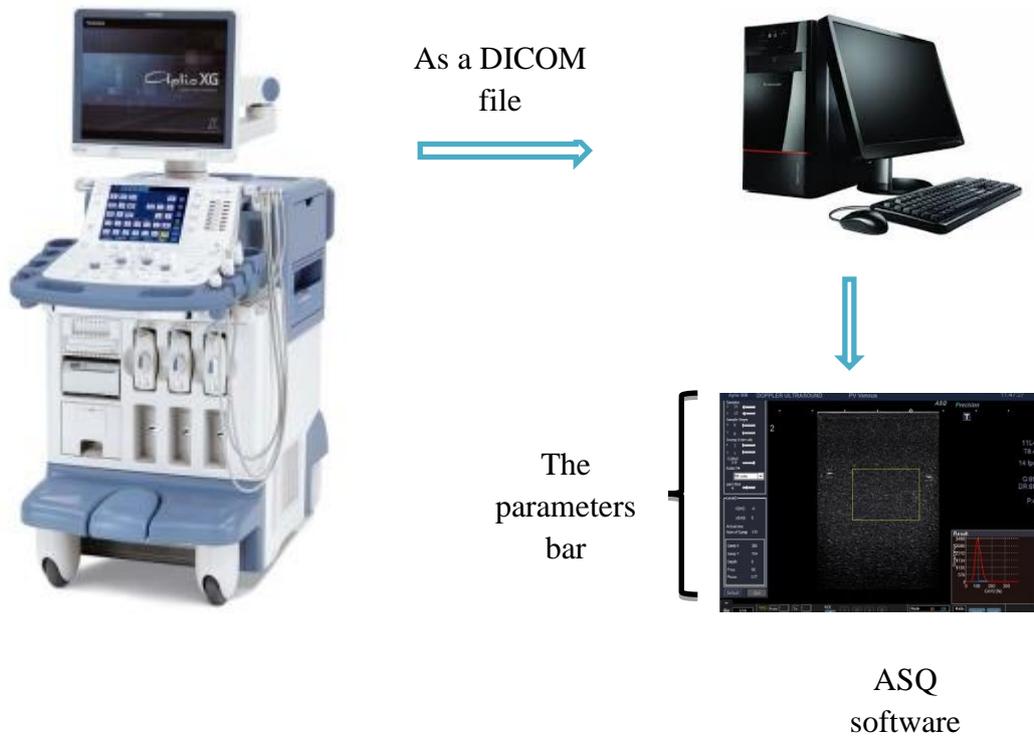


Figure 33: Block diagram of ASQ parameter assessment.

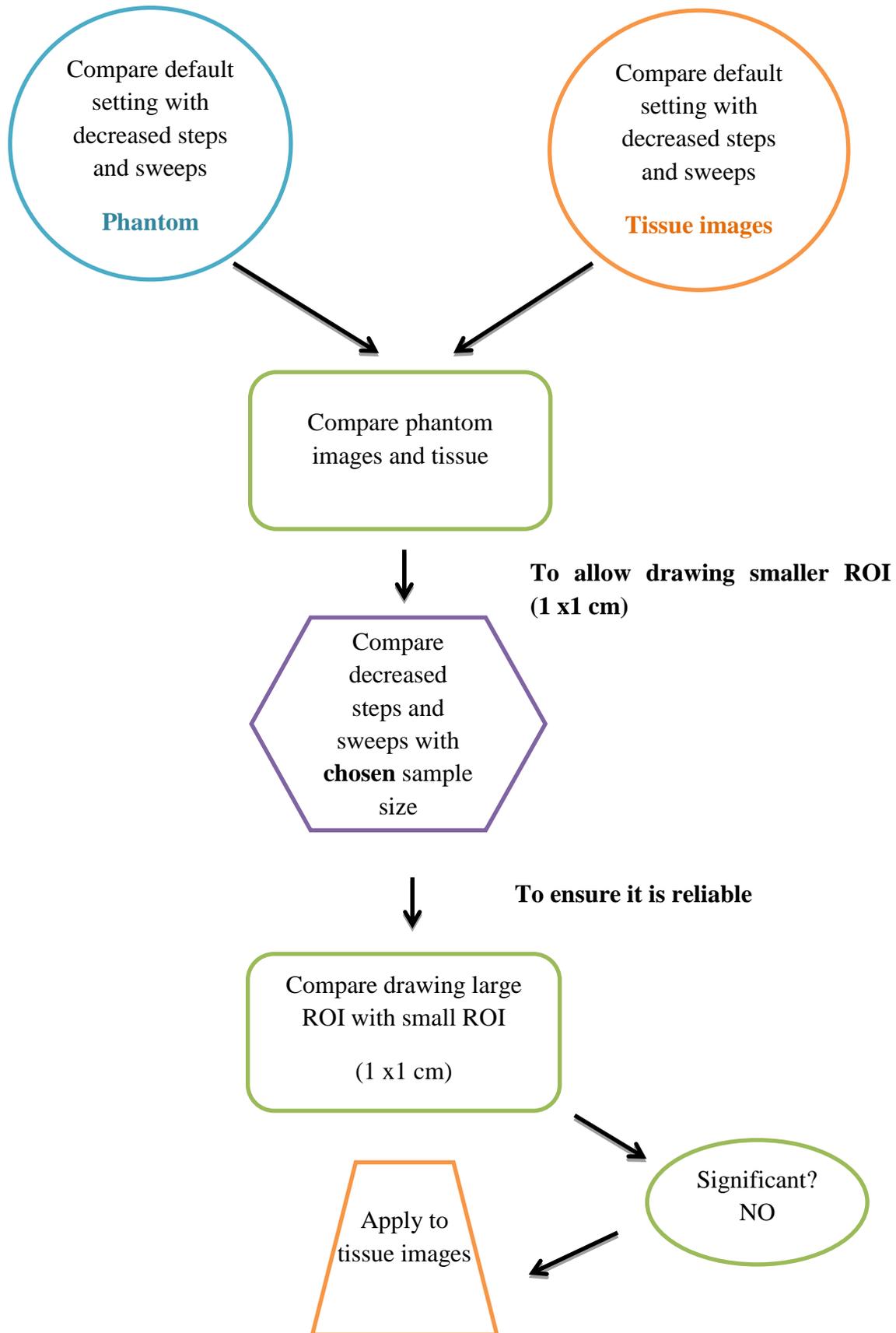


Figure 34: Flow chart of experiments

3.4.2. Results

When altering the x and y values of the sample, the values of (x 11 and y 13) were found to give an actual size of 115 samples (see Figure 35) and simultaneously allow us to draw the smallest ROI required (1 x 1 cm): please refer to Table 30 for a summary of the different values with their corresponding actual number of samples.

Sample values	Mean	Number of samples
x 17 y 19	101	255
x 15 y 17	101	207
x 13 y 15	99	151
x 11 y 13	98	115
x 9 y 11	96	79
x 7 y 9	91	51

Table 30: Summary of different x and y samples with their corresponding numbers of samples.

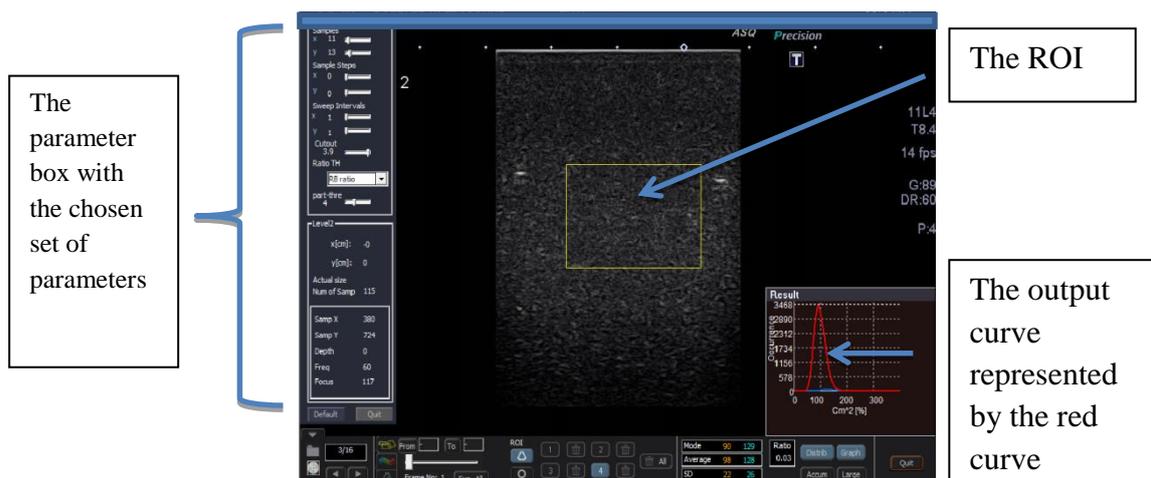


Figure 35: ASQ window with the chosen set of parameters.

After defining the best set of parameters, it was tested against tissue images (inhomogeneous tissue) to test for significance. When comparing the default setting with the decrease in sample steps and sweep interval to minimum, there was a significant difference between them, with a p-value of .004 (see Table 31). Similarly, when applying the decreased sample steps and sweep intervals setting to human tissue there was a significance difference, with a p-value of 0.043 (see Table 32).

Subsequently, the phantom images (homogenous) were compared to tissue images (inhomogeneous: see Figure 36) and found to have a significant difference between the means, with p-values of .001 in both the default and the decreased setting, which are summarised in Table 33.

	Median	P-VALUE	Significant?
Default setting	103		
Decrease steps 0,0 Sweeps 1,1	101	0.004	Yes

Table 31: ASQ Phantom images (homogeneous)

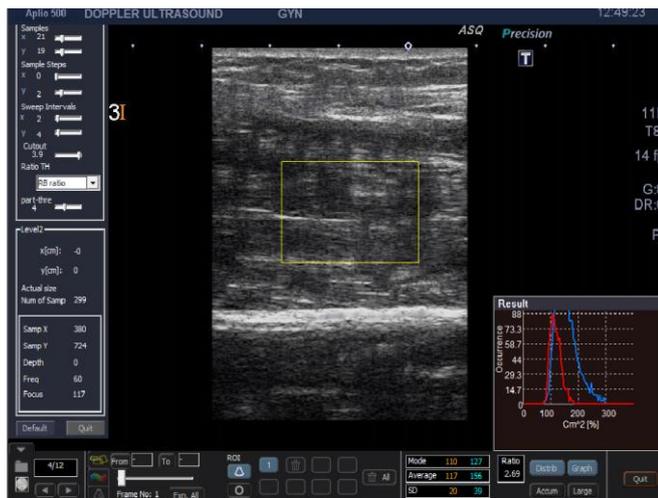


Figure 36: ASQ window for tissue image with default setting.

	Default setting	Decrease steps 0,0 sweeps 1,1	P-value
	Median	Median	
Red	119.5	113	0.043
Blue	163.5	167	0.912

Table 32: ASQ tissue images (inhomogeneous)

	Homogeneous	Inhomogeneous	P-value
	Median	Median	
Default	103	119.5	<.001
Steps0,0 sweeps 1,1	101	113	<.001

Table 33: The significance of the difference between the phantom images and the tissue images.

When using the pre-defined default settings, which are sample size of 299 with x 21 and y 19, the equipment will not allow the operator to draw a small ROI: hence, we had to find out which sample values of x and y would allow the drawing of small ROI (1 x 1 cm) without decreasing the sample size below 100, as advised by Toshiba engineers. All possible x and y values were tested to reach the required values. The resulting sample values were x 11 and y 13, which gives a sample size of 115. Afterwards, the chosen values were compared to the default sample size in tissue to test for significance. The p-value was .001, indicating a significant difference between the default sample size and the chosen one (please refer to Table 34).

	(x21,y19)	(x11,y13)	
Setting	Sample size 299	Sample size 115	P-value
	Steps 0,0 sweeps	Steps 0,0 sweeps	
	1,1	1,1	
	Median	Median	
Red	113	105.5	<.001
blue	167	172	0.529

Table 34: The effect of the sample size parameter on the mean of the curve in tissue images.

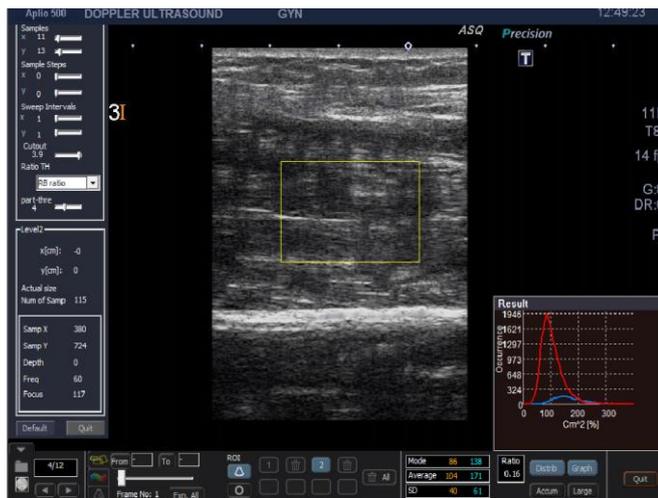


Figure 37: ASQ window for tissue image with chosen pre-set parameters.

To ensure that the new chosen parameters are reliable, we compared drawing large ROI to the desired small ROI (1 x 1 cm). There was no significant difference in the mean when using a small ROI or a large ROI with a p-value of .143 (see Table 35). Further testing was done on tissue images by applying the newly chosen parameters and comparing the large ROI with the anticipated small ROI (see Figure 37). No significant difference was seen, with a p-value of .165 (shown in Table 36).

Setting: (x11,y13)			
Sample size 115	Large ROI	Small ROI	P-value
Steps 0,0 sweeps 1,1		(1 x1 cm)	
Median	97	99	.143

Table 35: Comparison between drawing large ROI and smallest ROI with the new chosen setting on phantom images.

Setting (x11,y13)			
Sample size 115	Large ROI	Small ROI	P-value
Steps 0,0 sweeps 1,1		(1 x1 cm)	
	Median	Median	
Red	105.5	96.5	.165
blue	172	170	0.739

Table 36: Comparison between drawing large ROI and smallest ROI with the new chosen setting on tissue images.

3.4.3. Discussion

ASQ is a new, advanced technique that was introduced by Toshiba Medical Systems in 2009 to give a considerably higher resolution of ultrasound images. It has been proven, by several studies in the past few years, to be useful in diagnosing certain liver diseases (Onodera, 2013, Ricci et al., 2013, Wang et al., 2013). However, to apply this new method to other parts or organs of the body, such as the ovaries, it is essential to test it, as different organs have different normal textures and hence give different speckle patterns.

Testing the parameters of ASQ images was not an easy task. The reason for this is that very little information is available about this evolving application. Therefore, improvisation was needed to acquire the best results possible from the ultrasound images.

As seen from the results section, a set of parameters were chosen to be used on ovarian masses that allowed the drawing of small ROI (1 x 1 cm) without compromising the sample size. The reason for this is that the ovaries are much smaller in size than the liver: therefore, a smaller ROI was needed. No differences were observed when drawing small and large ROI, which ensured that the chosen parameters were reliable. When comparing the mean of phantom images to tissue images, a significant difference was observed as expected. This proves that the newly chosen set of parameters can be useful on inhomogeneous tissue and is not only applicable to phantoms.

3.5. (D) Applying ASQ to Images of Benign and Malignant Masses

The previous three sections explored ASQ and tested it for robustness as well as reliability and repeatability. They also tested the factors that could affect the output curve of the ASQ, and determined the best predefined parameters to be used. Consequently, we want to apply this new method of analysis to pelvic masses in an attempt to distinguish benign from malignant tissue. The aim of this section is to conduct a small study in which we apply ASQ analysis to images of pelvic masses to document the ability of ASQ in discriminating benign from malignant ovarian masses.

3.5.1. Methodology and Statistical analysis

To get the ethical approval, a full protocol for the proposed study of image texture analysis of Transvaginal ultrasound in diagnosing ovarian cancer was submitted to the Dyfed Powys Research Ethics Committee. The study received full ethical approval on 18th July 2013 (Ref. 13/WA/0206), as shown in appendix III. This project was also submitted to Cardiff and Vale NHS trust research and development committee, and was approved for commencement on 26th April 2013, as shown in appendix II.

Toshiba Aplio 500 was used in this study after it had been tested for reliability and repeatability (see section A in this chapter). However, only the Transvaginal probe (6 MHz) was used to acquire the images of pelvic masses.

ASQ software, PC-ASQR Version 1.11R001, was used to draw the different ROI sizes on the images and acquire the output curve. The statistical analysis (Q-Q plot, Shapiro Wilk test and Mann-Whitney U test) was performed using SPSS 17.0 for windows (SPSS Inc, Chicago, Illinois, USA).

3.5.1.1. Participants

Study population: women age 18 and above with known pelvic masses.

Inclusion criteria: All patients with confirmed pelvic masses and scheduled for surgery. Exclusion criteria: Patients with other Gynaecological malignancy, i.e. not pelvic masses, pregnant patients, patients with a previous history of bilateral oophorectomy, difficult scans and unclear scan images and age less than 18 years old.

Any patient with normal morphology or incomplete/ missing ultrasound results data were excluded from this study.

Patient withdrawal during the study: participants were free to withdraw from the study at any time.

3.5.1.2. Procedure: Recruitment and Consent

Participants were selected from a pelvic mass clinic and Gynaecology oncology clinics. All patients with known adnexal mass and scheduled for surgery were identified by members of the research team. The researcher approached the potential participant and informed them of the research study.

The study was explained and written information provided. Patients were given up to 24 hours to decide if they wanted to participate in the study, and then, after answering any question they had regarding the study, an appointment was booked for the scan.

On the day of the scan, written consent was obtained before starting the scan. All participants were booked for an ultrasound scan of the pelvis prior to their surgery.

The scanning procedure was the same for all participants in this study. A transvaginal transducer was used to perform the scan, and then, once the pelvic mass was identified on the screen, the ASQ button was enabled and the pre-set was on the ASQ setting. Images were stored as raw store clips on the machine hard drive. After the patients had left, images were transferred as DICOM files to a personal computer for off-line analysis. Then ASQ software was used to download the images. A ROI was drawn on the mass, which gives a graph that represents the normal and abnormal tissue in the ROI. The readings of the red and blue curves as well as the ratio between them were documented for each mass.

3.5.2. Results

A total of 45 masses were collected from 44 participants for the purpose of this study. Twenty-seven participants were premenopausal and seventeen were postmenopausal. Their ages ranged from 26 to 53 years in the premenopausal group (mean 41, median 43) and from 52 to 90 years (mean 64, median 58) in the postmenopausal group.

Histological specimens in all masses (n=45) included thirty-four benign masses, such as serous cystadenoma, mucinous cystadenoma, cystadenofibroma, dermoid, fibroid and endometrioma, and eleven malignant masses.

The output mean curves were categorised into two groups according to the type of lesion: benign or malignant. A Q-Q plot was used to determine the normality of the data by visual inspection. For normally distributed data, the plot should show a linear relationship; furthermore, the Shapiro-Wilk test was also used to test for normality. Figure 38 shows an example of normally distributed data (red curve - technique 2) while Figure 39 shows an example of non-normally distributed (blue curve - technique 1).

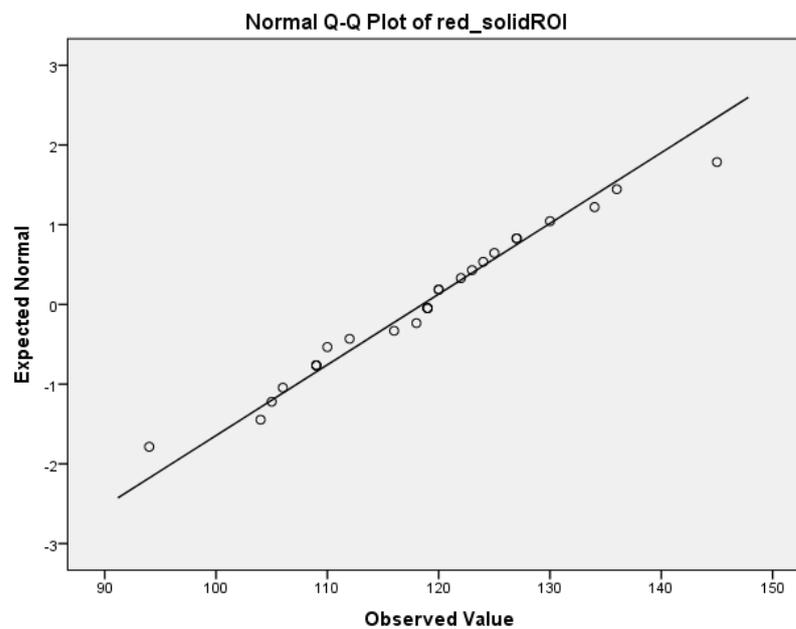


Figure 38: Example of Q-Q plot for normally distributed data (Red curve- technique 2)

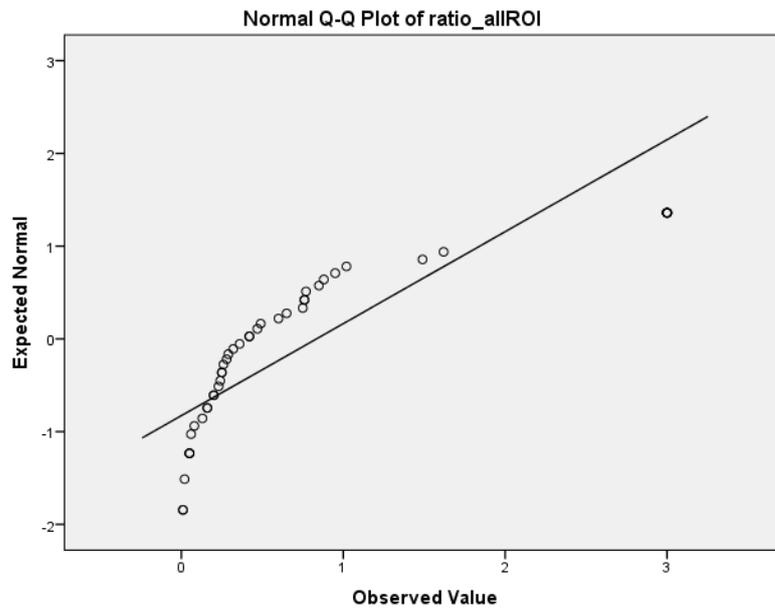


Figure 39: Example of Q-Q plot for a non-Normally distributed data (the ratio between the red and blue curve-technique 1)

	Shapiro-Wilk Sig.	Normally distributed?
Red curve (all mass included)	.119	Yes
Blue curve (all mass included)	.874	Yes
Ratio (all mass included)	<.05	No
Red curve (only solid area)	.963	Yes
Blue curve (only solid area)	.993	Yes
Ratio (only solid area)	<.05	No

Table 37: Shapiro-Wilk test results for normality of the data.

Two analysis techniques were used. The first included the whole mass in the ROI for analysis, while the second technique only included the solid area of the mass in the ROI.

The results of the Shapiro-Wilk test for normality are summarised in Table 37. They show that the red and blue curves both have normally distributed data in both techniques used (when $>.05$ is considered normal data), while the ratio between red and blue curves shows non-normally distributed data values $<.05$ in both techniques.

The mean and standard deviation of both the red and the blue curve and the ratio between them were calculated and are summarised in Tables 38 and 39 for the two different techniques respectively.

Type of curve	Mass type	Mean	SD
Red curve with all mass in ROI	Benign	123	30.8
	Malignant	123.5	14.4
Blue curve with all mass in ROI	Benign	175	39
	Malignant	184	24.4
The ratio between Red and Blue	Benign	.77	.99
	Malignant	1.02	1.08

Table 38: Summary of mean and standard deviation of the curves (ROI including the whole mass)

Type of curve	Mass type	Mean	SD
Red curve with only solid area in ROI	Benign	117.5	12.9
	Malignant	120.6	7.5
Blue curve with only solid area in ROI	Benign	168	22
	Malignant	177	25
Ratio between Red and Blue	Benign	.65	.9
	Malignant	.55	.4

Table 39: Summary of mean and standard deviation of the curves (ROI including the solid area only)

In the first technique, the means were similar in the benign and malignant masses, but they were different in the blue curve, with the malignant masses having the higher mean (184). In the ratio between the red and the blue curve, the malignant masses had higher mean (1.02) than the benign masses (.77).

In the second technique there were no similarities in mean values in all three groups (red, blue and ratio). Malignant masses had higher mean in the red and blue curve but lower mean in the ratio (.55) when compared to the benign masses.

Type of curve	Mass type	P-value	Significant different?
Red curve with all mass in ROI	Benign-Malignant	.956	NO
Blue curve with all mass in ROI	Benign-Malignant	.471	NO
The ratio between Red and Blue	Benign-Malignant	.489	NO

Table 40: Summary results of the significance test (ROI includes the whole mass)

Type of curve	Mass type	P-value	Significant different?
Red curve with only solid in ROI	Benign-Malignant	.518	NO
Blue curve with only solid in ROI	Benign-Malignant	.323	NO
The ratio between Red and Blue	Benign-Malignant	.748	NO

Table 41: Summary results of the significance test (ROI includes only the solid area).

The non-parametric Mann-Whitney U test was used to calculate the significance between the benign and malignant masses on data from both techniques. The results

when including the whole mass in the ROI show that the two groups were not significantly different in red, blue and the ratio between red and blue, with p-values > 0.05.

When using the second technique, which only included the solid area of the mass in the ROI, there were still no significance difference between the benign and the malignant masses in any of the curves, with p-values of .51, .32 and .74 for the red curve, the blue curve and the ratio between them respectively. The results for both techniques used are summarised in Tables 40 and 41.

An attempt to analyse a simple cyst was performed; however, there were no curves produced when applying the ROI on a simple cyst or any anechoic area. Therefore, the images of a simple cyst were not included in the analysis.

No further statistical tests were performed due to the non-significant differences between the two groups, with all p-values >.05.

3.5.3. Discussion and Conclusion

Acoustic structural quantification (ASQ) was originally designed by Toshiba Medical Systems to help in the diagnosis of liver diseases. It has been demonstrated to have the ability to separate homogenous from inhomogeneous tissue. In this section, an attempt was made to apply the ASQ analysis to pelvic masses in order to assess the

diagnosis of the nature of the mass by defining the difference between homogeneity of benign and malignant tissue.

Tissue characterisation can be used in conjunction with a visual interpretation of an ultrasound image to improve diagnosis. The aim of this section was to determine the ability of ASQ analysis in characterising ovarian tissue. In particular, we investigated the appropriateness of this method in discriminating benign from malignant tissue.

At the beginning of the study, the calculation of sample size was based on the assumption that ASQ may distinguish between benign and malignant masses with 5% significance level and 80% power: thus, a total of 200 women (100 participants with benign masses and 100 with malignant) were required for the study to demonstrate a significant difference. However, recruitment of participants for this study was somewhat difficult and slow. This was because we were dealing with women who had a suspicious mass and were waiting for surgery to have it removed: they were all very anxious, some were very ill and most of them did not want to have an additional internal scan, which can be uncomfortable in some cases when repeated in short period of time. Over a period of 18 months, only 44 women agreed to participate in this study, and after reviewing the preliminary results, a decision was made to stop the recruitment due to non-significant results.

It was agreed that there was no point in increasing the sample size if the results were not showing any discrimination ability between benign and malignant masses.

It is unknown whether a statistically significant difference between any of the parameters would have been found if it was practically possible to recruit a greater number of patients.

There were variations observed in all of the output curves values extracted from the same tissue type. For example, the SD of the extracted values from the red curve for benign tissue was 30.8 for technique 1 and 12.9 for technique 2. Similarly, the SD for the blue curve was 39 and 22 for technique 1 and technique 2 respectively.

In this study, two techniques for drawing the ROI were adopted. It can be seen from the results that selecting only the solid area from the mass did not improve the discrimination between benign and malignant masses. This is confirmed by the significance test results, whereby for technique 2, all output curve values were not statistically significant despite the fact that the mean values were different in both the benign and malignant groups (117.5 and 120.8 for the red curve and 168 and 177 for the blue curve). The results indicate that excluding the anechoic area from the analysis does not improve the diagnosis: this could be explained by the fact that the anechoic area originally does not give output curve values when analysed alone because only the images with raw data were used in the ASQ software.

This failure to discriminate the benign from the malignant masses using ASQ could be because most of the masses in both groups were fibrotic in nature, such as cystadenofibroma, fibroids, and adenocarcinoma.

This study has shown that quantitative ASQ analysis of B-mode images demonstrates an insignificant difference between benign and malignant tissue. This means that ASQ does not work on pelvic masses because both benign and malignant masses exhibit homogeneity and heterogeneity in the same way.

3.6. Summary

In this chapter, a new technique called ASQ was explored in four main sections. Section A provided background information about this new software and how it works, followed by a repeatability and reproducibility study, while section B tested the influence of four main factors on the ASQ output. Section C determined the pre-defined parameters of the ASQ software that can be used on ovarian tissue without compromising the image quality. Finally, section D explained the study that was done by analysing pelvic masses images using ASQ analysis.

3.6.1. Section A

A brief explanation of how ASQ works was introduced at the beginning of this chapter. The consistency of the image texture is an important factor in ASQ. The repeatability and reproducibility issues were therefore assessed to determine the consistency in the image texture, to which random variations inherent in the ultrasound scanner and the scanning process could contribute. This was done by calculating the CoV. A tissue-equivalent phantom was used for this purpose. This section also evaluated the agreement in the repeatability produced by two different operators. The results for the ASQ software repeatability show very low variation in the images produced under identical conditions, and good agreement was found in the images obtained by two different operators for all three probes used.

Based on finding from this section, it can be concluded that ASQ software is able to produce consistent images. This section has also established that intra-operator repeatability and good agreement between two operators are achievable. Therefore, it

is worth continuing to a similar assessment in a clinical setting, which involves more complex scanning processes: this is done in the next section.

3.6.2 Section B

The experimental evaluation of three transducers presented in this section aims to determine the robustness of ASQ software. A tissue-equivalent phantom was used as a test object. Several key factors that may affect the performance of ASQ were considered, namely ROI size, ROI depth, focal position, gain setting and transducer frequency. These factors were evaluated using p-values to determine significance. The ROI size was evaluated using two different sizes: large and small. The ROI depth dependency was evaluated by drawing ROI at two different depths: 2 cm and 4 cm. The focus was positioned at 2 cm and then at 4 cm to test for significance on the mean of the output curve. The gain was set at 100 % and 85 % to evaluate its effect on the mean of the output curve. All available frequencies in each transducer were tested. Three frequencies were found for the linear probe (8.4, 7.2 and 6.2 MHz) while four frequency setting were available for the curved probe (6, 5, 4 and 3 MHz). In the TV probe, three frequencies were available for testing (8, 7.2 and 6 MHz).

Of the four factors that might influence the ASQ output, focal position was found to significantly affect the output in the images from the TV probe. On the other hand, the linear probe was significantly dependant on three of the factors, namely ROI size, ROI depth and focus position. Interestingly, the curve transducer was independent from all factors. In other words, the mean of the output curve was not affected by any of the four factors in the images acquired by the curved probe. Transducer frequency

did not affect the mean of the output curve in all three transducers except when using high frequencies such as 8 MHz.

The findings from this section permit further study in a clinical setting with the TV probe with confidence that the output measurements are reliable, with caution when using focal position.

3.6.3 Section C

The pre-defined parameters of the ASQ window were tested in this section to determine the appropriate parameters values to be used on ovarian masses. (Please refer to Figure 27 for the flow chart of the experiment). Since the default setting was provided originally with the software to be used on liver tissue, it was necessary to find out the best parameters to be used with ovaries. The values of x 11, y 13, sample steps of 0, 0 and sweep intervals of 1, 1 with an actual sample size of 115, were found to be suitable for the application of ASQ on ovarian masses. Further testing of a smaller ROI size (1 x 1 cm) was done on both phantom and tissue images to test for significance. There were no significant differences between larger ROI and the desired smaller ROI in both phantom and tissue images, which permit the use of those chosen parameters in the clinical setting and their application to smaller masses.

The phantom images were then compared to tissue images and demonstrated to have significant differences between the means, which shows that these sets of parameters are applicable to human tissue and not only valid on phantom images.

3.6.4. Section D

In this section, a prospective study was conducted by applying ASQ to images of pelvic masses. An explanation of the recruitment procedure was given, followed by the results of the study. Forty-five masses were saved as raw images to be analysed using ASQ software. Data were tested for normality using both the Shapiro-Wilk test and Q-Q plot. Then when non-normally distributed data were identified, a non-parametric statistical test was used (Mann-Whitney U test) to calculate the p-value for the significance of the difference. Lastly, a discussion and conclusion to the study was provided, concluding that ASQ cannot differentiate between the homogeneity of benign and malignant tissues.

4. Texture Analysis Features

This chapter will investigate the possible use of texture analysis features in the diagnosis of pelvic masses by applying these features to ultrasound images of the masses. The chapter is divided into three main sections. Section A will explain the materials and methods used in this research; then section B will show the results produced after collecting and analyzing the data, and section C will discuss these results and compare them to other studies done in the same field.

The aim of this chapter is to investigate the use of texture analysis features, namely the Grey Level Co-occurrence Matrix (GLCM) and the Wavelet feature, as previous research Hamid (2011) has concluded that they are the best features to be used to diagnose ovarian masses.

In this previous preliminary study, the author investigated the robustness of five different texture analysis features using phantom. It was found that GLCM and wavelet were not affected by factors such as, ROI size, ROI depth and gain setting. After that the chosen texture analysis features were applied to thirty ultrasound images (ten normal ovaries, nine cysts and eleven malignant masses). A significant difference was found between all categories with 91% sensitivity between normal tissue and malignant and between cysts and malignant using GLCM and 89, 90% using wavelet feature, respectively (Hamid, 2011). However, Hamid did not consider comparing the malignant masses with benign masses.

Therefore, this current study will use a larger number of prospective participants and compare benign and malignant masses using GLCM and wavelet as well as comparing the benign subgroups such as: endometrioma, fibroids and dermoids to malignant masses. After that the results will be compared to other widely used scoring systems such as, RMI, PMI and ADNEX, which has not been done before in any other study. As well as dividing the data into pre- and postmenopausal groups and analyse them separately.

(A) Materials and Methods

The materials used in this study were introduced and explained in the previous chapter, which provided an overview of the equipment and the principles of Ultrasound and transvaginal transducers (Please refer to Chapter 3, section A).

In this chapter, texture analysis concepts and texture analysis software (MaZda 4.6, Institute of Electronics, Technical University of Lodz, Poland) are described. After that the methods used and the procedures followed to accomplish this study, including the study design, sample size and recruitment of participants, are explained.

4.1. Materials

4.1.1. MaZda Software

In order to extract the textural features from the ultrasound images, MaZda software version 4.6 (Institute of Electronics, University of Lodz, Poland) was used in this study. A brief explanation for this software is given in this section. More details about this software are available online at: <http://www.eletel.p.lodz.pl/mazda/>

MaZda is an abbreviation derived from the term ‘Macierz Zdarzen’, which means ‘co-occurrence matrix’ in Polish. It was developed in 1996 specifically for the analysis of image texture (Szczypiński et al., 2009). It was the first program created to conduct a quantitative analysis of texture within a freely selected region of interest (ROI) and to provide an interpretation of computed results. At first, the use of MaZda was aimed at analysis of MRI images. Nevertheless, it revealed its effectiveness in analysis of other types of images such as x-ray, CT and ultrasound images (please refer to Figure 40).

This program is the most widely used for texture analysis. There are only few other examples of image texture analysis of software available. The other non-commercial packages, such as KeyRes and LS2W, provide only a limited set of the functions that are present in MaZda (Szczypiński et al., 2009).

MaZda is used for texture analysis in 2D and 3D images. It provides a complete path for quantitative analysis of image textures, including computation of texture features, procedures for feature selection and extraction, data classification methods, and various data visualization and image segmentation tools.

The MaZda package is a reliable and robust tool for the analysis of image textures, as confirmed by (Szczypiński et al., 2009).

4.1.2. Region of interest (ROI) selection

Region of interest are sets of pixels in 2D images selected to be analysed. In order to avoid unnecessary processing, defining a specific region of interest (ROI) will concentrate the computation effort on the part of the image that is under investigation.

This is supported by a study done in 2003 about an automatic technique for

morphological analysis, which found that better performance was seen when ROIs were obtained manually. This was explained by the fact that using a fully automatic algorithm would involve an additional error factor (the process of automatically finding the margins of the mass): (Zimmer et al., 2003).

MaZda software ROIs can be of arbitrary shape. It allows up to sixteen ROIs within a single image. These regions can be overlapped if needed. In this study, only one ROI is drawn around the mass and used in the analysis, with some cases of drawing several ROIs around the cystic areas to be excluded from the analysis of the complex mass, as performed by (Hamid et al., 2011).

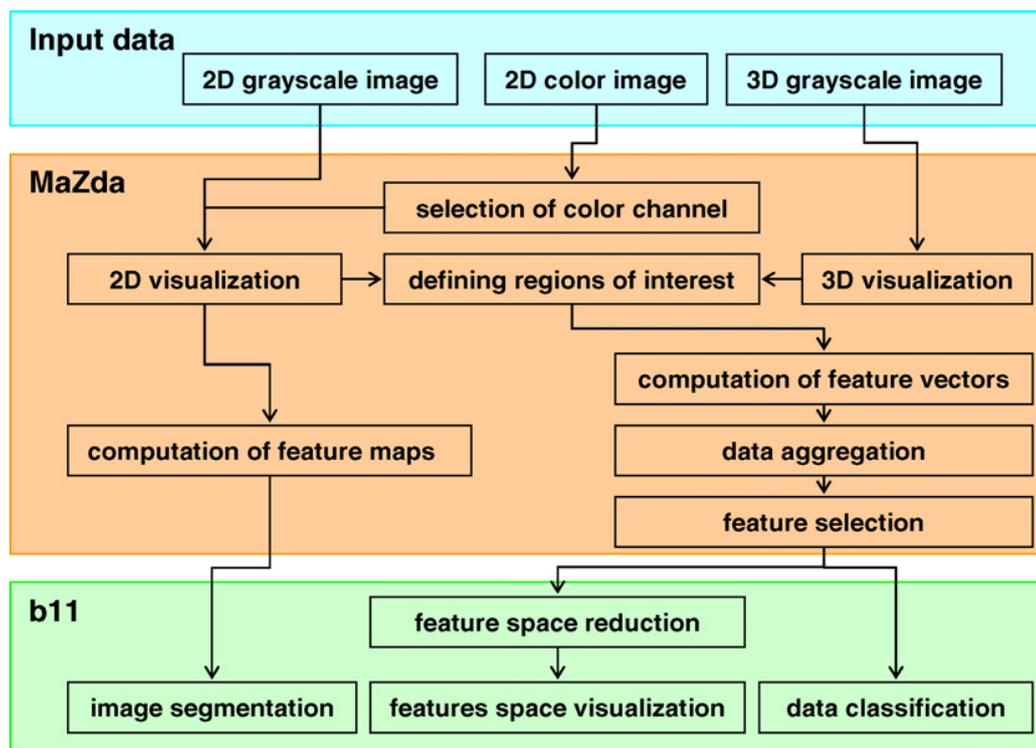


Figure 40: flowchart of analysis pathways in Mazda package (adopted from (Szczyński et al., 2009))

4.1.3. Texture features

MaZda software computes six types of textural features: Histogram-based (first-order statistics), GLCM, RLM, Gradient, AR and Wavelet. Based on previous preliminary research (Hamid et al., 2011) only two of the six features, which are GLCM and Wavelet, were used in this study. Hamid and colleagues explained that the histogram-based features were not included in their study because they are computed from the intensity of pixels without considering any spatial relations between pixels within the image and supported their choice using findings from Bader et al. (2000) that showed insignificant results when using a histogram-based texture feature to distinguish benign tumours from carcinomas.

4.1.3.1. Grey level Co-occurrence matrix (GLCM)

Harliack (1979) was the primary author who proposed the co-occurrence matrix parameters. GLCM is a second-order statistical technique that allows for the extraction of statistical information from the image regarding the distribution of pairs of pixels. It is computed by defining a direction, a distance and a pair of pixels separated by this distance, computed across the defined direction, which are analysed. (M. Hajek et al., 2006, Alqahtani et al., 2010, Singh et al., 1997, Beekman and Visser, 2004). Numerous co-occurrence matrices can be computed for a single image, one for each pair of pixels with a defined distance and direction.

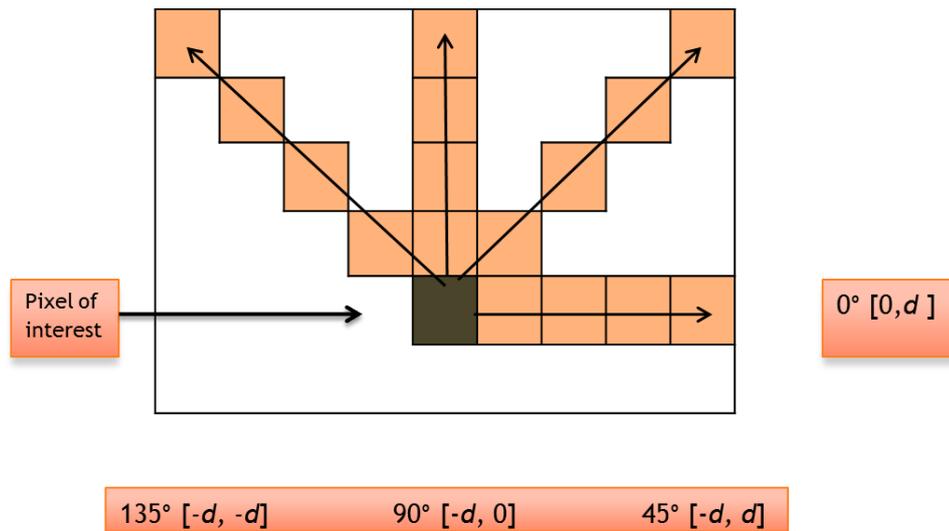


Figure 41: Showing computation of co-occurrence matrix with pixel of interest (Lu et al., 2006)

Normally a set of 220 co-occurrence matrices are computed, for five different pixels distances from 1 to 5 pixels, and four different directions ($\Theta=0^\circ, 45^\circ, 90^\circ$ and 135°) (see Figure 41). Please refer to section 2.1.2.6 in Chapter 2 for background information.

4.1.3.2. Wavelet

This is a transform method of texture analysis. The principle of transform methods is that the texture properties are represented in space where the coordinate system is closely related to the feature of texture, such as frequency and size (Materka A and M, 1998). The wavelet images are scaled up to five times in both horizontal and vertical directions: this results in image transformation into twenty frequency channels. Energies computed within the channels provide data on texture frequency components

and are used as texture characterizing features (Szczypiński et al., 2009, Castellano et al., 2004).

4.1.3.3. Tissue characterisation

It is evident from the available literature that visual inspection alone for the diagnosis of tissue pathologies can be prone to bias and can lack sensitivity and reliability. One main reason for this is that diagnosis largely depends on the sonographer, who observes tissue characteristics from the image and compares them with images of different pathologies to come out with a diagnosis. The other alternative to this method is the application of invasive methodologies, such as needle-guided biopsy. Although this is a very robust technique with less chance of error, it has the disadvantage of being invasive in nature and is therefore too impractical to be the method of choice for all patients.

To address this problem, researchers have developed quantitative criteria with the help of computer systems to aid diagnosis. The use of texture analysis described in this chapter attempts to address this shortcoming of subjective analysis and provide a better diagnostic tool to distinguish benign from malignant ovarian masses. As described in the section above, the commonly used texture analysis methods are based on grey level distribution of individual pixels and their relationship with adjacent pixels.

4.1.4. Data acquisition and image analysis

The role of texture analysis of medical images has become increasingly important in the field of diagnostic imaging. Many researchers have focused on assisting texture analysis through the development of computer-aided diagnosis systems. These systems have the ability to enhance the precision and accuracy of image characterisation.

A computer-aided quantitative evaluation system for adnexal masses may be divided into five parts:

1. Acquiring an image: a device such as MRI, CT or ultrasound can be used to generate the digital image. In this study, an ultrasound machine was used.
2. Defining the region of interest (ROI) within the image. A different size and shape of the ROI can be defined within an image as well as multiple ROI per image.
3. Extraction of the texture features from the defined ROI in the image.
4. Determination of the most suitable texture parameters to analyse the image for the purpose required.
5. Data analysis and classification of the texture features within the image.

4.2. Method

In this section, subject recruitment is described, an overview of patient management is considered and the statistical methods used in data entry and analysis are described.

4.2.1. Study Design

This study was a quantitative prospective cross-sectional study on women with known pelvic masses who were booked for surgery.

4.2.1.1. Choice of Methods

4.2.1.1.1. Sample size

In setting up the study, it was important to recruit a sufficient number of participants to demonstrate the differences between the two groups.

Statistical advice was obtained to determine the required number for each arm, in order to test the null hypothesis (the primary end point): i.e. the ability of GLCM and wavelet features to analyse images of ovarian masses and categorise them as benign or malignant. However, there have been no previous studies on texture analysis to distinguish between benign and malignant ovarian masses. Moreover, studies which used other morphological scores have suffered from limitations, including the subjectivity of the score, because it depends on the sonographer's experience. In view of these factors, the calculation of sample size was based on the assumption that texture analysis may distinguish between benign and malignant masses with 5% significance level and 80% power: thus, a total of 200 women (100 participants in each arm) were required for the study to demonstrate a significant difference.

The use of power calculation provides a scientific basis for the number of subjects required to make up the sample size that is needed to reject the null hypothesis with a given level of significance (usually 80%).

4.2.2. Participants

Study population: women aged eighteen and above with known pelvic masses.

Inclusion criteria: All patients with confirmed pelvic masses and scheduled for surgery.

Exclusion criteria: Patients with other Gynaecological malignancy, i.e. not pelvic masses, pregnant patients, patients with a previous history of bilateral oophorectomy, difficult scans and unclear scan images and patients aged less than eighteen years old.

Postmenopausal status, defined as amenorrhea of twelve months following natural or surgical menopause, or one year or greater of hormone-replacement therapy begun for menopausal symptoms (Menon et al., 2008, 2009). In the current study, only women with confirmed adnexal masses were included. Any patient with normal morphology or incomplete/missing ultrasound data was excluded from this study.

Patient withdrawal during the study: participants were at liberty to withdraw from the study at any time.

4.2.3. Procedure: Recruitment and Consent

Participants were selected from a pelvic mass clinic and Gynaecology oncology clinics. All patients with known adnexal mass were identified by members of the research team. The researcher approached the potential participants and informed them of the research study.

The study was explained and written information was provided. Participant information sheets and consent forms were provided so that the patients could find out more about the study before deciding whether or not to participate. (Please see Appendices IV and V). Patients were given up to 24 hours to decide if they wanted to participate in the study, and then, after answering any questions they had regarding the study, an appointment was booked for the scan.

On the day of the scan, written consent was obtained before starting the scan. All participants were booked for an ultrasound scan of the pelvis prior to their surgery.

Randomization/ treatment assignment: the researcher was blinded from any ultrasound scan reports done for the participants in clinics to avoid bias in the diagnosis of ovarian lesions.

Side effect: Ultrasound is a safe modality with no documented side effects. Transvaginal scanning is a routine procedure that does not raise any problems except slight discomfort by some patients. To handle any other side effects that might arise, a qualified healthcare professional was present at all times during scanning.

4.2.4. Demographic Data

The following demographic data were obtained for all study women: age, menopausal status, history of previous hysterectomy.

Postmenopausal status was defined as the absence of menses for a minimum of twelve months or an age of more than fifty years in women who had a prior hysterectomy.

Family history was considered positive if the patient had a first-degree relative (i.e. mother, sister, or daughter) or a second-degree relative (i.e. grandmother, granddaughter, aunt or niece) with documented ovarian or breast cancer.

4.2.5. Ultrasound

The scanning procedure was the same in all participants in this study. All ultrasound procedures were undertaken by the researcher or the co-researchers only for the purpose of this study.

Ovarian morphology dimensions and volume were reviewed. The volume was determined using the formula for an ovoid ($d1 \times d2 \times d3 \times 0.532$).

Ovarian morphology was classified on ultrasound as:

Normal: - ovary of uniform hypoechogenicity and with a smooth outline with or without a single inclusion cyst or spots of calcifications

- Inclusion cyst must be single, less than 10 mm in diameter and not distorting the outline of the ovary;
- Simple cyst: a single, thin-walled, anechoic cyst with no septa or papillary projections;
- Complex: any case in which the ovary has any non-uniform ovarian echogenicity, excluding single simple or inclusion cysts. (Menon et al., 2009);

In women with bilateral ovarian masses, data from both sides were used for the analysis, i.e. complex morphology and simple cyst or dermoid on one side and a complex mass on the other one.

The following morphological ultrasound information was recorded in each case: volume of the ovary, site and volume of the cyst, cystic wall structure, and cystic wall thickness, presence of septation and septal thickness, presence of solid areas within the cyst, papillation height if present and echogenicity.

Presence or absence of Doppler signal and the site of the signal were documented for each cyst.

In addition to applying texture analysis features (GLCM and wavelet), RMI, PMI and ADNEX scoring systems were calculated for the masses in this study for comparison of the diagnostic performance. Please refer to Chapter 2 for a brief explanation and tables of scores of the RMI, PMI and the ADNEX model.

Histopathological diagnosis was obtained in women who underwent surgery and used as the gold standard. In cases where participants were managed conservatively and no histology results were available, ultrasound diagnosis by an expert examiner was used in typical adnexal masses (endometrioma, typical dermoid and simple cyst) as well as the use of a second diagnostic model such as MRI or CT where appropriate. In addition, for those with no histology, follow-up at a minimum of twelve months after the ultrasound scan was used as recommended in the ROCKETS study protocol. This project is a new ongoing study that aims to develop and validate new risk prediction

models that estimate the probability of having ovarian cancer in women with symptoms (unpublished) (Saundar et al., 2015).

4.2.6. Data Analysis

All data were statistically analysed using the Statistical Package for Social Sciences (SPSS) program version 17.0 for windows (SPSS Inc, Chicago, Illinois, USA), Database Access 2010 and Microsoft Excel 2010 software in order to calculate the statistical differences and measure the variances. This software is widely used instrument in statistiscs and social sciences. Therefore, the researcher considers the two software applications as appropriate to be used for this piece of research.

Assumed Sensitivity 97% (95% CI 91.5% to 99.0%) based on 100 malignant.

Assumed Specificity 85% (95% CI 76.7% to 90.7%) based on 100 benign.

Obtaining these figures would serve as proof of the principle that the new method is capable of performing at least as well as existing methods, and it has the advantage of being objective, with minimal operator dependence anticipated. These assumptions were calculated by a professional medical statistician prior to commencing the study.

As long as the data is normally distributed, the use of parametric tests is the appropriate method to test independent data (Lang and Secic, 2006b).

The nonparametric statistical technique known as chi-square was used to examine the categorical demographic variables. The Mann-Whitney U test was used in this study, as it is often viewed as the nonparametric correspondent of the Student t-test and can

be used where the data is suspected to be not normally distributed. Moreover, the Mann-Whitney test is less likely than the t-test to falsely indicate significance if outliers are present (Motulsky and Brown, 2006). Continuous demographic variables were compared using independent t-tests if data were normally distributed and non-parametric tests when the data were not normally distributed.

Ninety-five percent confidence intervals were calculated where appropriate. The alpha level was set at 0.05 and any p-value less than the alpha level is considered statistically significant. This is a conventionally used level to ascertain differences and confirm significance between groups, which therefore, by definition, is statistically significant (Lang and Secic, 2006b).

In statistics, logistic regression is used to estimate the possibility that a specific event will occur by using a number of predictor variables which might be either numerical or categorical and fitting the data to a logit function (logistic curve) (Kleinbaum, 2010). Logistic regression has a down-side of overestimating the Odds ratio if the sample size is less than 500. By increasing the sample size, this overestimation contracts until it reaches the true population value. However, in a single study, overestimation due to the small sample size might not affect the interpretation of the results due to the fact that this will be much lower than the standard error of the estimate (Kleinbaum, 2010). In this study, logistic regression was used at the end of the analysis to explore a different way of analysing the data by using the variables in a logistic regression model.

4.2.6.1. Performance measures

Sensitivity, specificity, positive predictive value (PPV), and accuracy were calculated to evaluate the performance of the texture features. The true negative (TN) is the number of benign samples identified as benign. The true positive (TP) is the number of malignant images identified as malignant. The number of malignant masses detected as benign is quantified by the false-negative (FN) measure. False-positive (FP) is the number of benign samples identified as malignant.

The sensitivity is the proportion of actual positive (malignant) cases which are correctly identified, while the specificity is the proportion of actual negative (benign) cases which are correctly identified. The PPV is the ratio of true positives to combined true and false positives, and the accuracy is the ratio of the number of correctly classified samples to the total number of samples. (Please refer to Table 42 for the calculation of the performance tests).

PPV	$\frac{TP}{TP + FP}$	Sensitivity	$\frac{TP}{TP + FN}$
NPV	$\frac{TN}{TN + FN}$	Specificity	$\frac{TN}{TN + FP}$
Accuracy	$\frac{TP + TN}{TP + TN + FN + FP}$		

Table 42: Calculation of statistical performance tests. (TP is true positive, FP is false positive, TN is true negative and FN is false negative)

4.2.7. Location of Study and Access Arrangements

This study took place in the Medical Physics and Clinical Engineering Department and Doppler Department at the University Hospital of Wales (UHW), Cardiff, UK. As well as Pelvic mass clinic in UHW . No other departments or laboratories were included.

4.2.8. Ethical Considerations

This study uses human participants: therefore, there are several important ethical issues relating to the design and implementation of the study. The ethical principle governing research is that patients should not be harmed as a result of participating in the research (Silverman, 1985). Moreover, the dignity, rights, safety and well-being of participants in research must be of primary consideration (Power and Kuyken, 1998).

The ethical committee provides independent expert opinion on whether the proposed research is ethical and respects the dignity, rights, safety and well-being of participants. For this reason, a full protocol for the proposed study of image texture analysis of Transvaginal ultrasound in diagnosing ovarian cancer was submitted to the Dyfed Powys Research Ethics Committee. The study received full ethical approval on 18th July 2013 (Ref. 13/WA/0206), as shown in appendix III. This project was also submitted to Cardiff and Vale NHS trust research and development committee, and was approved for commencement on 26th April 2013, as shown in appendix II.

4.3. Results: texture analysis

A total of 226 patients fitted the criteria and were recruited for this study in the period between November 2013 and May 2015. Of these patients, 63 (27.8%) were then excluded from the analysis for several reasons, such as participants refusing to have the internal scan (n=20), incomplete data required for the analysis (n=7), no histology found (n=34) three months after the scan due to participants being managed conservatively and not proceeding to surgery, and finally participants agreeing to have the scan and then cancelling (n=2).

Therefore, only 163 patients (169 masses) were analysed for the purpose of this study, of whom 95 were premenopausal and 68 postmenopausal. Their ages ranged from 18 to 53 years in the premenopausal group (mean 41, median 42) and from 52 to 91 years (mean 67, median 66) in the postmenopausal group. In total, 169 ovarian masses were used for texture analysis.

Histopathology results of the masses showed 116 benign, 29 malignant and 24 simple cysts.

A total of 81 of the benign masses were premenopausal and 35 postmenopausal. Of the malignant masses, 8 were premenopausal and 21 postmenopausal and of the simple cyst, 9 were premenopausal and 15 postmenopausal.

For the purpose of the analysis, the benign group were further divided into subgroups using ultrasound diagnosis, including 32 mature teratomas (dermoid), 27 endometriomas, 24 fibroids and 33 other suspicious benign masses. The histology

results of these suspicious masses were documented and are summarised in Table 43. The most commonly identified of these suspicious benign masses were the fibroma (18%) and both the serous and mucinous cystadenoma (15%) as well as the cyst adenofibroma (15%).

The malignant tumours were classified using histopathology and the results are summarised in Table 44. They include both serous and mucinous adenocarcinoma, adult granulosa cells, poorly differentiated adenocarcinoma and borderline serous and mucinous cystadenoma and a couple of unknown malignant classification. The most commonly found type was the borderline serous cystadenoma (24%) followed by the high grade serous adenocarcinoma (17%) and the grade 3 carcinoma (14%).

Type of mass	Frequency	Percentage (%)
Fibroma	6	18.2
Serous cystadenoma	5	15.2
Cyst adenofibroma	5	15.2
Adenofibroma	2	6
Mucinous cystadenoma	5	15.2
Leiomyoma	2	6
Epithelial cysts	2	6
Adenomyosis	1	3
Endometriosis	1	3
Degenerated fibroid	1	3
Pedunculated fibroid	1	3
Benign unknown type	2	6
Total	33	100%

Table 43: Summary of histology results of 33 suspicious benign masses.

Type of mass	Frequency	Percentage (%)
High grade serous adenocarcinoma	5	17.2%
Invasive squamous cell carcinoma	1	3.5%
Adult granulosa cell	2	6.9%
Poorly differentiated adenocarcinoma	2	6.9%
Appendice mucinous neoplasm	1	3.5%
Endometrial adenocarcinoma	1	3.5%
Grade 3 carcinoma	4	13.8%
Borderline serous cystadenoma	7	24%
Borderline mucinous cystadenoma	3	10.3%
Borderline serous epithelia neoplasm	1	3.5%
Unknown classification	2	6.9%
Total	29	100%

Table 44: Summary of histology results of 29 malignant masses

For each group (cyst, benign and malignant masses), the mean, median, standard deviation and standard error of the mean (s.e.) of the extracted features were calculated and summarised in Table 45. The mean values of the GLCM and wavelet features are presented visually in Figure 42 and Figure 43, respectively.

Of the three groups, the malignant tissue exhibited the highest GLCM mean value (512), while cysts exhibited the lowest GLCM value (90). The mean value for the benign tissue is 320. The same pattern was observed in the wavelet feature, with the malignant tissue having the highest mean value (30809) and the cyst having the lowest mean value (3823).

Texture feature	Tissue type	Mean	Median	SD	S.E.
GLCM	Cyst	90	60	82	18
	Benign	320	250	229	21
	Malignant	512	376	341	63
Wavelet	Cyst	3823	7714	4360	930
	Benign	21642	17191	15286	1407
	Malignant	30809	26337	21036	3906

Table 45: Summary of the mean, median, standard deviation, and standard error of the mean of the extracted features.

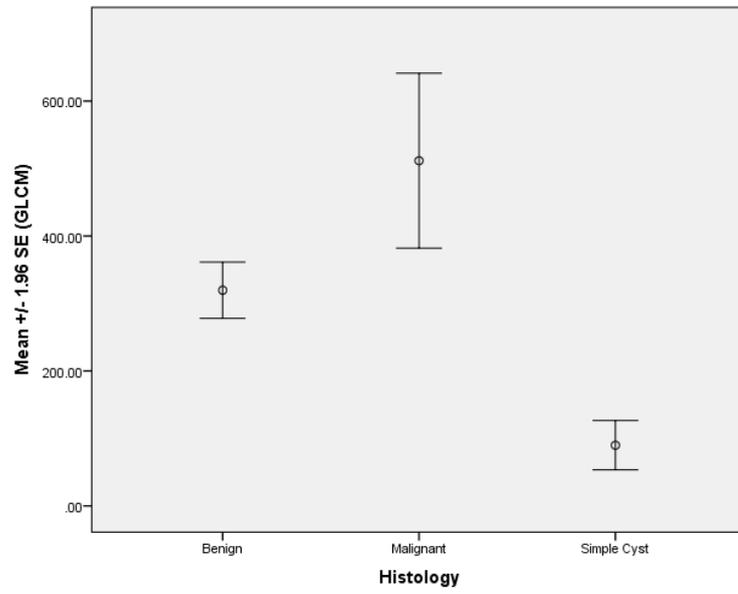


Figure 42: Graph showing mean value of GLCM feature for benign, malignant and simple cyst tissue.

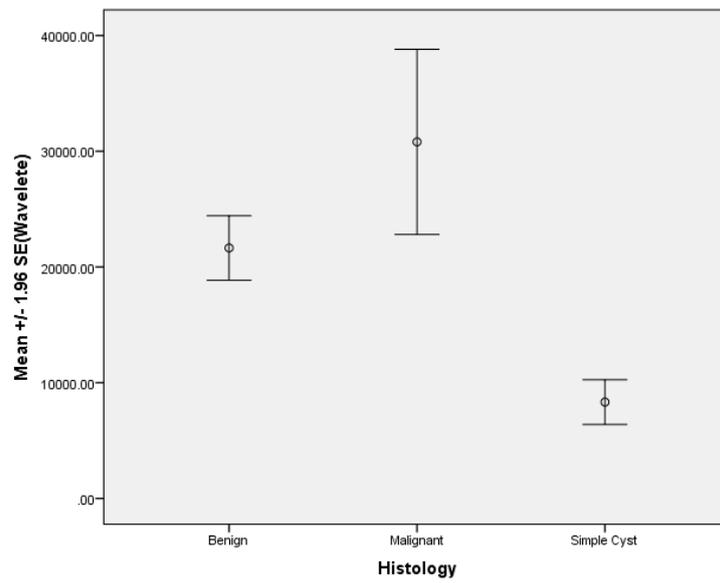


Figure 43: Graph showing mean values of wavelet feature for benign, malignant and simple cyst tissue.

4.3.1. Normality testing

Before testing the significance of the difference between means, it is important to know whether the data is normally distributed in order to decide which test to use: either parametric statistical tests such as t-tests for normally distributed data or the non-parametric tests such as the Mann-Whitney U test for non-normally distributed data.

The Shapiro-Wilk test was used to test for normality and it was found that most of the data in this study were not normally distributed, except for the wavelet feature in the cysts, which appeared to be normally distributed (.427). Results are summarised in Table 46.

Furthermore, a Q-Q plot was used to visually inspect the normality of the data. It was noticed that in the GLCM feature, not all tissue types were following the line, which means that they were not normally distributed. Please refer to Figures 44 to 49 for Q-Q plots of different types of tissue in both the GLCM feature and the wavelet feature.

Texture feature	Tissue type	Shapiro-Wilk test (sig.)	Normally distributed?
GLCM	Cyst	<.05	No
	Benign	<.05	No
	Malignant	.031	No
Wavelet	Cyst	.427	Yes
	Benign	<.05	No
	Malignant	.007	No

Table 46: Shapiro-Wilk test for normality

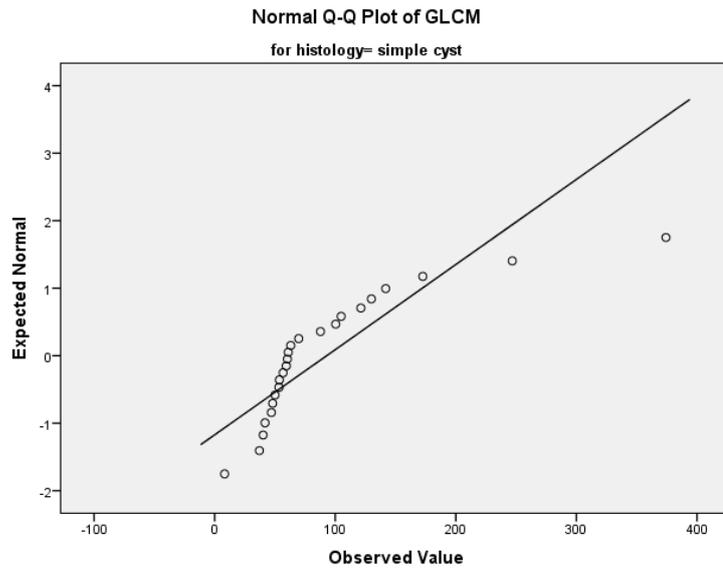


Figure 43: Q-Q plot for GLCM in the simple cyst tissue.

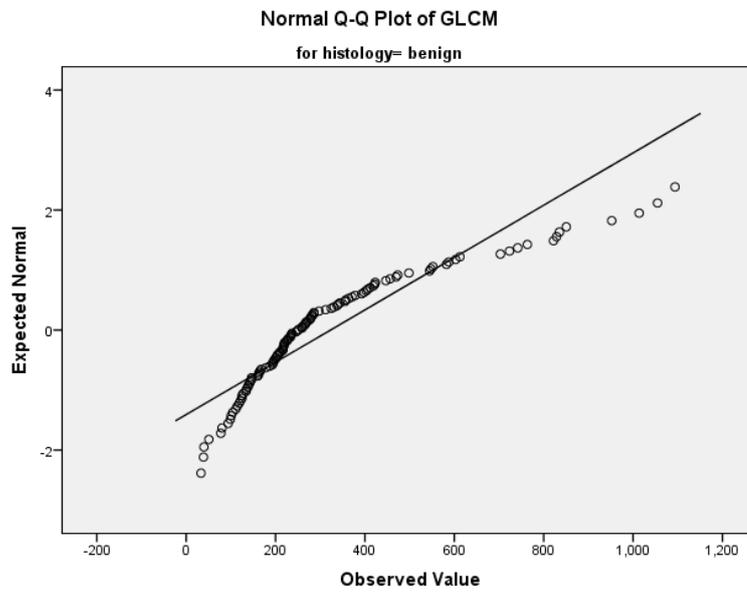


Figure 44: Q-Q plot for GLCM in the benign tissue.

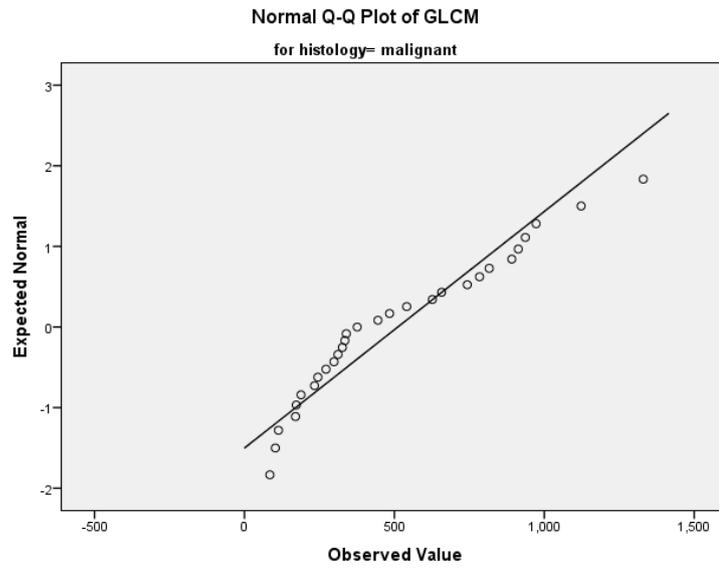


Figure 45: Q-Q plot for GLCM in the malignant tissue.

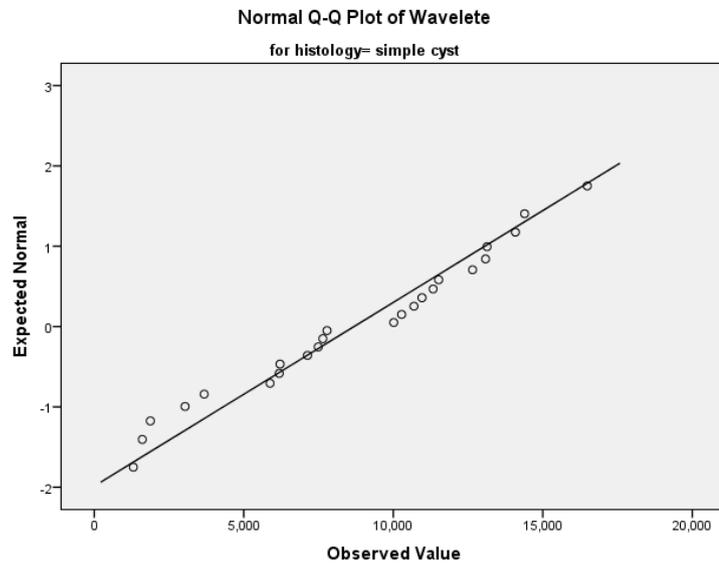


Figure 46: Q-Q plot for wavelet in the simple cyst tissue.

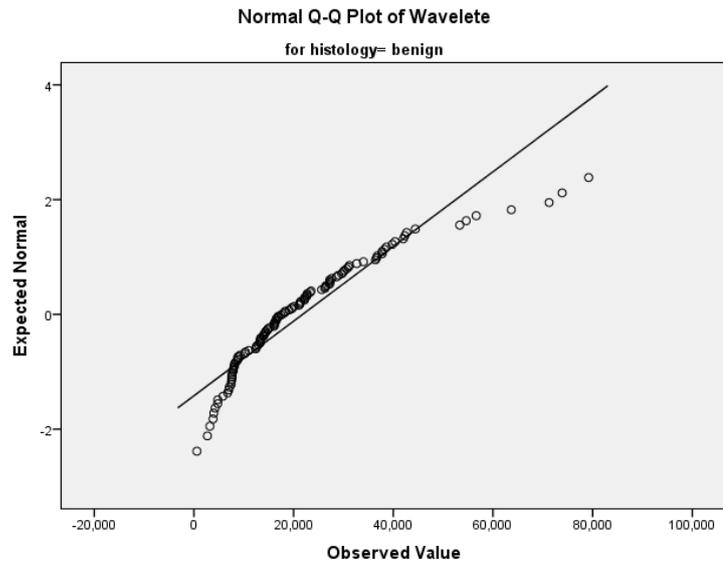


Figure 47: Q-Q plot for wavelet in the benign tissue.

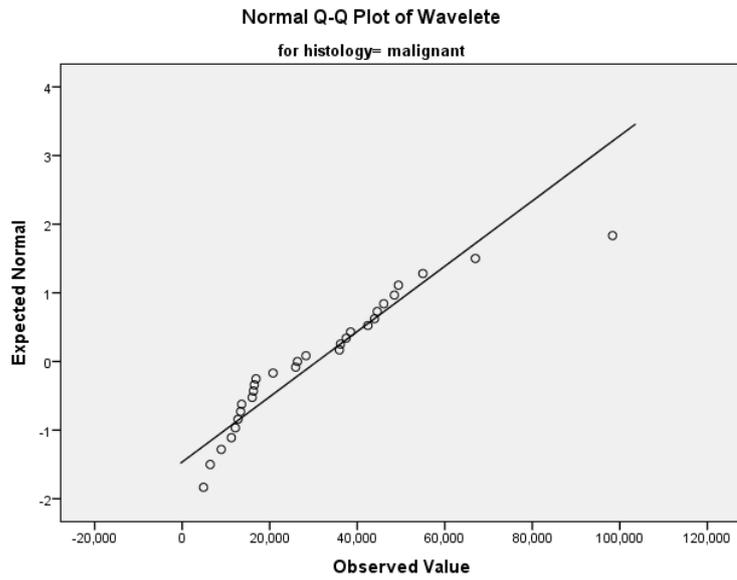


Figure 48: Q-Q plot for wavelet in the malignant tissue.

4.3.2. Test for significance

The values of the extracted features were compared in pairs: benign and malignant; cysts and malignant; and benign and cysts. The results for the GLCM feature show that all differences between group pairs were statistically significant: the p-value was $<.05$ in all groups.

The results for the wavelet feature were similar to the GLCM for all three groups, with a significant difference of $p <.05$ in all groups. The results for both GLCM and wavelet features are summarised in Table 47.

Texture feature	Group pair	P-value	Significantly different?
GLCM	Benign-Cyst	$<.05$	Yes
	Benign-Malignant	.004	Yes
	Malignant	$<.05$	Yes
Wavelet	Benign-Cyst	$<.05$	Yes
	Benign-Malignant	.027	Yes
	Cyst-Malignant	$<.05$	Yes

Table 47: Summary results of the significance test in all group pairs for both GLCM and wavelet.

For further analysis, the benign group was sub-divided into four groups, namely teratomas, endometriomas, fibroids and other suspicious or difficult to diagnose benign masses. When comparing these subgroups to the suspicious benign masses, it

was found that GLCM could significantly differentiate the benign suspicious group from the teratomas ($p=.021$) and from the endometriomas ($p=.011$). On the other hand, GLCM could not differentiate the benign masses from the fibroids ($p=.449$).

Texture feature	Group pair	P-value	Significantly different?
GLCM	Benign-teratoma	.021	Yes
	Benign-endometrioma	.011	Yes
	Benign-fibroid	.449	No
Wavelet	Benign-teratoma	.014	Yes
	Benign-endometrioma	.170	No
	Benign-fibroid	.693	No

Table 48: Tests of significant difference for the benign subgroups.

The wavelet feature only significantly differentiated between the benign suspicious group and the teratomas ($p=.014$); however, there was no significant difference between the benign suspicious group and the endometriomas or between the benign suspicious group and the fibroids ($p=.170$, $p=.693$) respectively. Results are summarised in Table 48.

Texture feature	Group pair	P-value	Significantly different?
GLCM	Malignant-teratoma	.697	No
	Malignant-endometrioma	<.05	Yes
	Malignant -fibroid	.009	Yes
Wavelet	Malignant -teratoma	.988	No
	Malignant -endometrioma	.004	Yes
	Malignant -fibroid	.061	No

Table 49: Tests of significant difference test for malignant vs. benign subgroups.

When comparing these subgroups of benign masses with the malignant masses, it was found that the GLCM feature could significantly differentiate between the malignant masses and both endometriomas ($p < .05$) and fibroids ($p = .009$), but it could not differentiate them from the teratoma masses ($p = .697$).

When using the wavelet feature to compare groups, a significant difference was noticed between the malignant masses and the endometriomas ($p = .004$), while in the other two groups, comparing the malignant masses with the teratomas ($p = .988$) and with the fibroids ($p = .061$), no significant differences were seen. These p-values are summarised in Table 49.

4.3.3. ROC and AUC

Receiver operating curve (ROC) analysis was performed to determine the ability of the GLCM feature to discriminate between cysts and benign masses, and between benign and malignant masses (i.e. for those group pairs that demonstrate a statistically significant difference ($p < 0.05$)). The corresponding ROC curves are shown in Figure 48 & 49 (between benign and malignant masses), Figure 50 & 51 (between cyst and benign masses), and Figure 52 & 53 (between cysts and malignant masses). An AUC close to 1 indicates a strong discriminatory power/ability of the indicator variable, while the AUC close to 0.5 indicates that the variable has little discriminatory power.

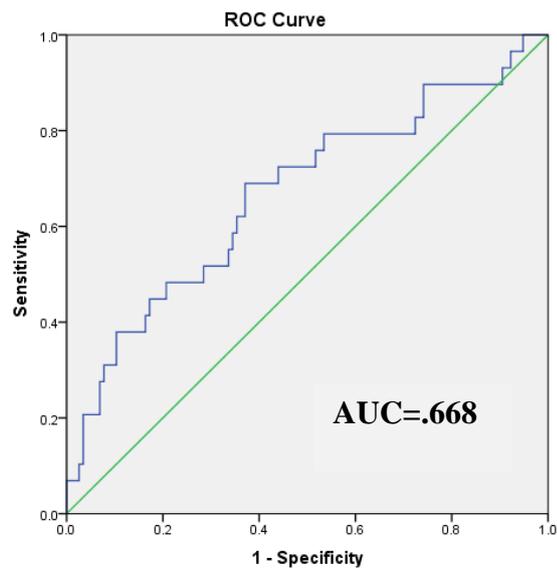


Figure 49: ROC curve for GLCM feature to discriminate between benign and malignant masses

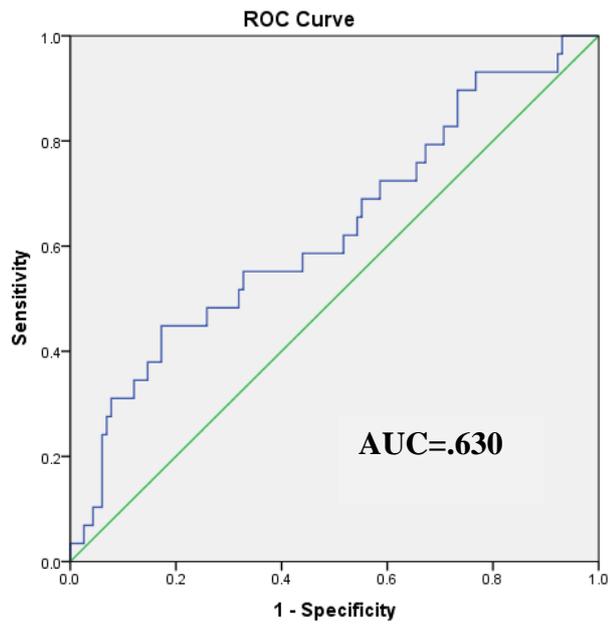


Figure 50: ROC curve for Wavelet feature to discriminate between benign and malignant masses

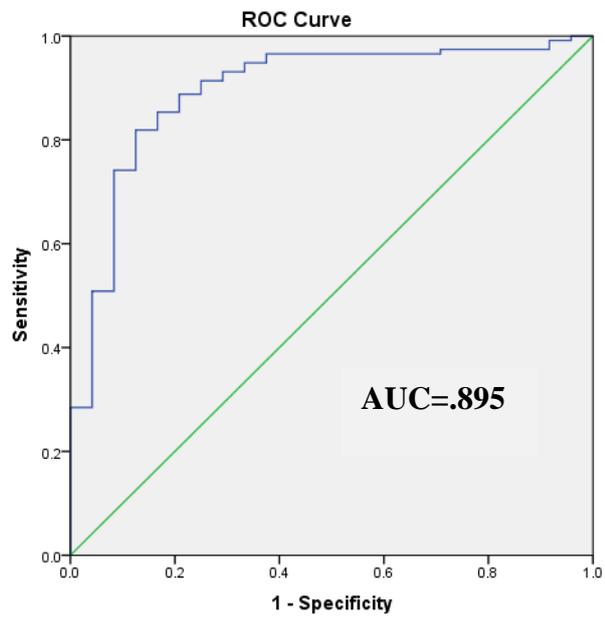


Figure 51: ROC curve for GLCM feature to discriminate between cysts and benign masses

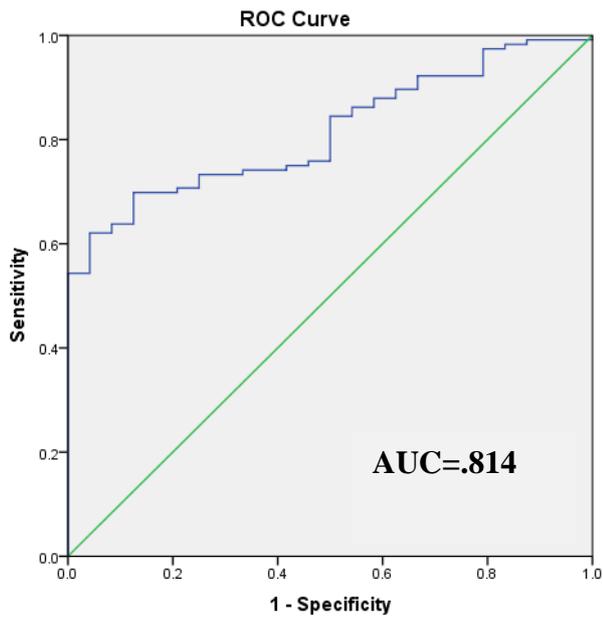


Figure 52: ROC curve for wavelet feature to discriminate between cysts and benign masses

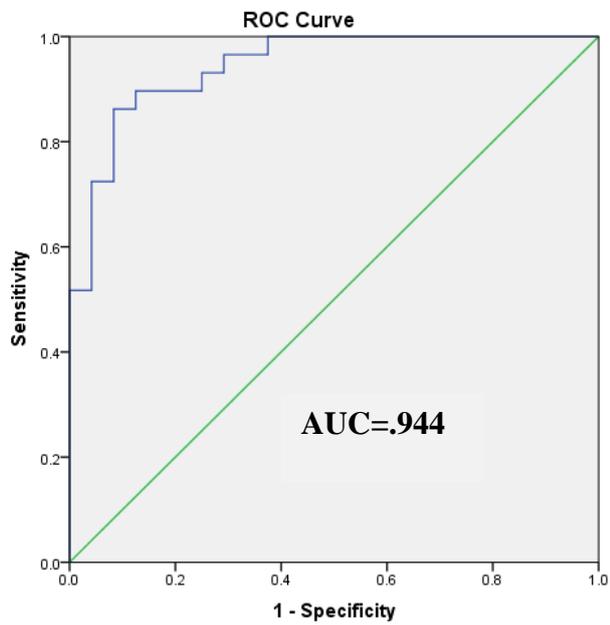


Figure 53: ROC curve for GLCM feature to discriminate between cysts and malignant masses

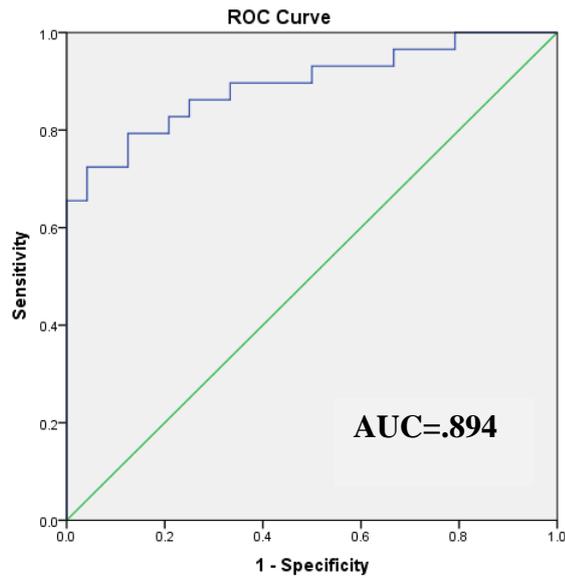


Figure 54: ROC curve for Wavelet feature to discriminate between cyst and malignant masses

Table 50 presents the area under the curve (AUC), which reflects the ability of the GLCM and the wavelet features in discriminating between the groups. It shows that the GLCM is a poor indicator (AUC=.668) for distinguishing between benign and malignant masses; on the other hand, it is an excellent indicator in both discriminating between malignant masses and cysts (AUC=.994) and between benign masses and cysts (AUC=.895). Results are similar when applying the wavelet feature, which is a poor indicator (AUC=.630) to differentiate between benign and malignant masses but a good indicator (AUC=.894) to differentiate between malignant masses and cysts as well as between benign masses and cysts (AUC=.814).

To evaluate the ability of the texture feature to correctly classify masses, a threshold value was selected to get the highest possible sensitivity and specificity. For instance, in discriminating between benign and malignant masses, the use of 245 as a threshold

value led to an estimated sensitivity of 72% and specificity of 60%. Table 51 presents examples of the threshold value in both GLCM and wavelet features.

Texture feature	Group pair	AUC	S.E.	Discriminatory ability
GLCM	Benign-Malignant	.668	.060	Poor
	Malignant-Cyst	.994	.029	Excellent
	Benign-Cyst	.895	.037	Excellent
Wavelet	Benign-Malignant	.630	.061	Poor
	Malignant –Cyst	.894	.044	Good
	Benign-Cyst	.814	.037	Good

Table 50: Area under the curve (AUC) with standard error (S.E.) associated with both GLCM and Wavelet features

Texture feature	Group pair	Threshold value	Sensitivity (%)	Specificity (%)
GLCM	Benign-malignant	245	72%	60%
	Malignant-Cyst	86	97%	62%
	Benign-cyst	100.5	93%	70%
Wavelet	Benign-malignant	17191	60%	60%
	Malignant-Cyst	10484	90%	59%
	Benign-cyst	10826	72%	62%

Table 51: Sensitivity and Specificity associated with GLCM and Wavelet features.

Furthermore, when comparing the malignant masses with the benign subgroups, ROC curves were generated only for the group pairs that showed a significant difference. Therefore, an ROC curve was created for the malignant–endometrioma pair and for the malignant–fibroid pair for the GLCM feature and for the malignant–endometrioma pair using the wavelet feature. Results showed that the GLCM feature had a good (AUC=0.8) discriminatory ability in distinguishing between malignant masses and endometriomas, while it presents a fair ability (AUC= .711) when distinguishing malignant masses from fibroids. Moreover, the wavelet feature revealed a fair ability to distinguish between malignant masses and endometriomas (AUC= .723). Results are summarised in Table 52. ROC curves for these subgroups are presented in Figures 54, 55 and 56.

Texture feature	Group pair	AUC	Discriminatory ability
GLCM	Malignant-endometrioma	.8	Good
	Malignant- fibroid	.711	Fair
Wavelet	Malignant-endometrioma	.723	Fair

Table 52: The area under the curve (AUC) associated with both GLCM and Wavelet features in the subgroups

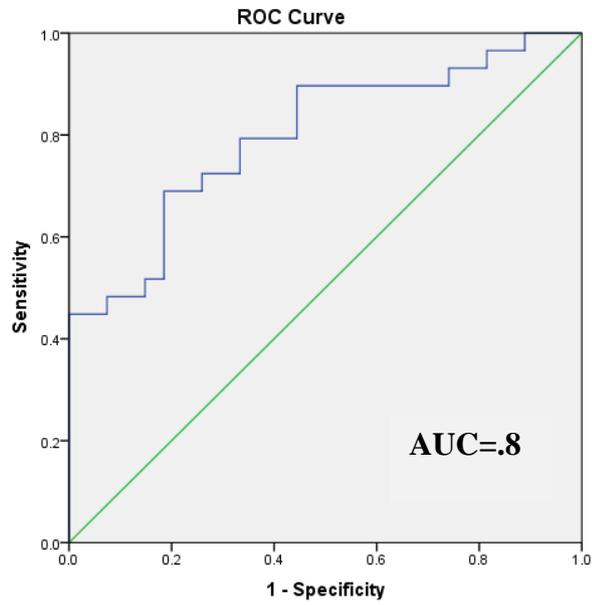


Figure 55: ROC curve of the GLCM feature to discriminate between malignant masses and endometriomas

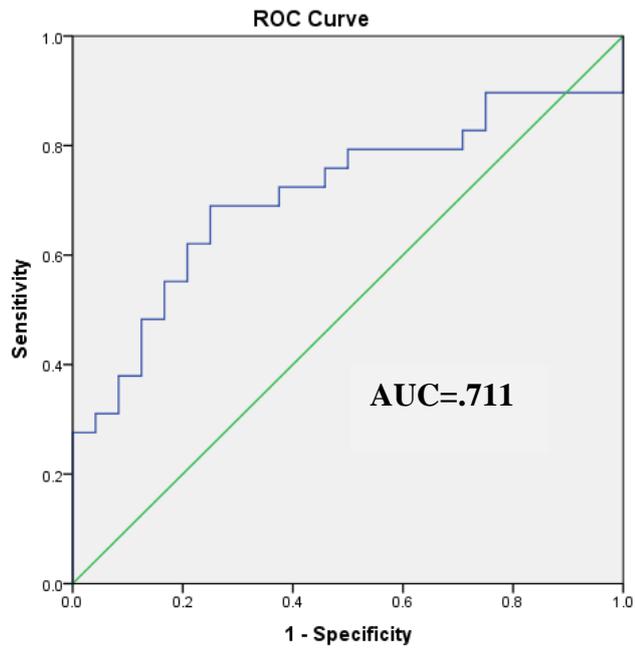


Figure 56: ROC curve of the GLCM feature to discriminate between malignant and fibroid masses

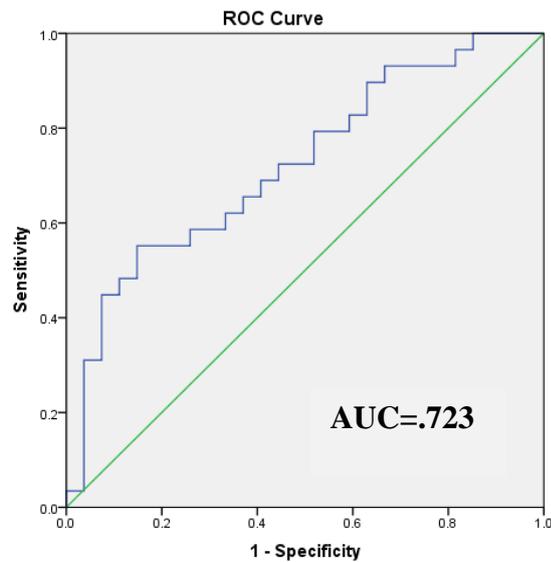


Figure 57: ROC curve of the wavelet feature to discriminate between malignant masses and endometriomas

To evaluate the ability of the texture feature to correctly classify masses in these subgroups, a threshold value was again selected here to get the highest possible sensitivity and specificity. However, sensitivity was prioritised over specificity. In the GLCM feature, a threshold value of 292.3 was found to give a sensitivity of 70% and a specificity of 82% when attempting to distinguish between malignant masses and endometriomas. For comparison between malignant masses and fibroids using the same texture feature, a threshold value of 265.3 was found to give a sensitivity of 72% and a specificity of 63%.

In the wavelet feature, a threshold value of 16242.5 was seen to give a sensitivity of 70% and a specificity of 60% when comparing malignant masses and endometriomas. Results are summarised in Table 53.

Texture feature	Group pair	Threshold value	Sensitivity (%)	Specificity (%)
GLCM	Malignant-endometrioma	292.3	70%	82%
	Malignant-fibroid	265.3	72%	63%
Wavelet	Malignant-endometrioma	16242.5	70%	60%

Table 53: Sensitivity and specificity associated with GLCM and Wavelet features in the subgroups

4.3.4. Diagnostic scoring systems

Further assessment was done using the widely used scoring systems, which are PMI on grey scale, RMI and the ADNEX model.

4.3.4.1. RMI

When applying RMI scores to the data, only 99 masses were applicable, of which 14 were excluded due to missing CA125 values. The remaining 85 masses were divided into three categories when applying the RMI scoring system: the high risk group with $RMI > 250$, the moderate group with RMI between 25-250 and the low risk group with $RMI < 25$. According to RMI, 15 (17.6%) masses were in the high risk group, 32 (37.6%) in the moderate risk group and 38 (44.7%) in the low risk group, of which 11 had a score of 0 due to a U score of 0. In the high risk group ($RMI > 250$), 8 masses were diagnosed as benign by histology, which means that these masses gave false positive results. On the other hand, 15 masses were diagnosed as malignant in the low risk group (< 250), which means that they were false negative results. Please refer to Table 54.

Type of the test	Malignant by histology	Benign by histology
RMI >250	True positive (TP)	False positive (FP)
	7	8
<250	False negative (FN)	True negative (TN)
	15	55
Total	22	63

Table 54: 2x2 contingency table for RMI scores

4.3.4.2. PMI

In this study, we also applied the PMI score to compare it with our texture analysis results. In total, 102 of 169 masses were eligible for the PMI score. The results were divided into three groups: low risk (-2 – 0) intermediate (1-2) and high risk (above 3). Forty (39.2%) masses were found to be low risk, while thirteen (12.7%) were in the intermediate group and forty-nine (44.1%) were in the high risk group. In the high risk group, thirty-six masses had been diagnosed as benign by histology, which means that they were false positive results, while in the low risk group, three masses were diagnosed as malignant in the histology, which means that they were in the false negative group. Please refer to Table 55.

Type of the test	Malignant by histology	Benign by histology
High risk and intermediate risk	(TP) 26	(FP) 36
Low risk	(FN) 3	(TN) 37
Total	29	73

Table 55: 2x2 contingency table for PMI score

4.3.4.3. ADNEX

The new model made available by the IOTA group – the ADNEX model – was applicable to 81 of the 169 masses in this study. Fifty-two (64%) masses were considered most probably benign and twenty-nine (36%) were considered to be at high risk of malignancy. In this high risk of malignancy group, eleven masses were benign according to histology, which makes them false positive results, while in the low risk group, another eleven masses were found to be malignant by histology, which categorises them as false negatives. Please refer to Table 56.

Type of the test	Malignant by histology	Benign by histology
High risk >50%	(TP) 18	(FP) 11
Low risk <50%	(FN) 11	(TN) 41
Total	29	52

Table 56: 2x2 contingency table for the ADNEX score.

4.3.4.4. Sensitivity and Specificity

The following table (Table 57) presents a summary of the diagnostic performance of the texture features of interest and the most commonly used scoring systems. PMI showed the best overall sensitivity (90%) and NPV (93%) amongst the indices; however, it had the lowest specificity of them all (51%). In contrast, RMI showed the highest specificity (87%), while compromising on sensitivity, which was the lowest in all indices (32%) and NPV (80%). GLCM and wavelets present similar specificity (60%) with somewhat higher sensitivity in the GLCM (72% compared to 60%) in the wavelet and NPV (90%). Moreover, the ADNEX score showed a similar sensitivity (62%) and NPV (80%) to the wavelet, with a much higher specificity (80%).

Correlation	Sensitivity	Specificity	PPV	NPV	Accuracy
GLCM	72% (95% CI: 53-91%)	60%	25%	90%	60%
Wavelet	60% (95% CI: 37-83%)	60%	22%	90%	60%
RMI	32% (95% CI: 0-65%)	87%	50%	80%	73%
PMI	90% (95% CI: 80-100%)	51%	42%	93%	62%
ADNEX	62% (95% CI: 29-95%)	80%	62%	80%	73%

Table 57: Summary of the sensitivity, specificity, PPV, NPV and accuracy of the texture analysis features GLCM and wavelet compared to RMI, PMI and ADNEX.

4.3.5. Tissue characterisation

In this study, a comparison was done between masses to test the ability of the selected texture feature to distinguish between different types of mass. In this section, we explore further by testing the ability of these features to distinguish between different

types of tissue. For this purpose, a sample of twenty normal ovarian tissues was analysed using both GLCM and wavelet. Then it was compared to the analysis of solid areas taken from malignant masses.

The results show a significant difference between the two different tissues with $p < .05$ when using the GLCM, with a mean value of 190.6 for the normal ovary compared to 503 in the malignant tissue and $p = .004$ when using the wavelet feature, with a mean of 17089 in the normal ovary and 38761.5 in the malignant tissue. Please refer to Table 58 for a summary of the results.

Texture Feature	Tissue type	Mean	Median	P-value	Significantly different?
GLCM	Normal ovary	190.6	167.3	<.05	Yes
	Malignant	503	464.5		
Wavelet	Normal ovary	17089	14254	.004	Yes
	Malignant	38761.5	31810		

Table 58: The significance test between normal ovarian tissue and malignant solid tissue.

4.3.6. Premenopausal group

Subdividing the study population into the categories of pre- and postmenopausal status allowed more in-depth analysis of the performance of the three indices.

Ninety-eight women were found in the premenopausal group, with eighty-one benign, eight malignant and nine simple cysts. Similar to the total population analysis, group

of pairs were compared to each other to test for significance. When using the GLCM, all group pairs were still found to have a significant difference in the premenopausal group. However, when applying the wavelet feature, the malignant and benign masses could not be differentiated significantly ($p=.366$). The other two groups of pairs (benign-cyst and malignant-cyst) still had significant differences. Results are illustrated in Table 59.

Texture Feature	Group pair	P-value	Significantly different?
GLCM	Benign-cyst	<.05	Yes
	Malignant-benign	.022	Yes
	Malignant-cyst	<.05	Yes
Wavelet	Benign-cyst	<.05	Yes
	Malignant-benign	0.366	No
	Malignant-cyst	<.05	Yes

Table 59: Summary of results of the significance test in premenopausal women

Following the analysis of the total population, ROC curves were generated and AUC were calculated. As mentioned above, only the group pairs with significant differences were used in these ROC curves. When distinguishing between benign and malignant masses, GLCM showed a fair ability, with $AUC=.747$ in the premenopausal group compared to the total population analysis, which had poor ability ($AUC=.668$). Please refer to Table 60.

In the other group pairs – the benign compared to the simple cysts – both GLCM and wavelet showed good discriminatory ability, with AUC=.879 and .850 respectively. In the group pair of malignant compared to simple cysts, both features showed excellent discriminatory ability in the premenopausal population (AUC= .944, .986) for GLCM and wavelet respectively. ROC curves for all pairs are illustrated in Figures 59 to 63. These results are similar to the total population analysis. Please refer to Tables 50 and 60 for the comparison.

Texture feature	Group pair	AUC	S.E.	Discriminatory ability
GLCM	Benign-malignant	.747	.079	Fair
	Benign-Cyst	.879	.070	Good
	Malignant-cyst	.944	.059	Excellent
Wavelet	Benign-malignant	.597	.105	Poor
	Benign-cyst	.850	.045	Good
	Malignant-cyst	.986	.022	Excellent

Table 60: The area under the curve (AUC) and standard error (S.E.) associated with both GLCM and Wavelet features in premenopausal women

Similar to the total population, a threshold value was selected to get the highest possible sensitivity and specificity. In the GLCM feature, a threshold value of 279.5 gave a sensitivity of 75% and a specificity of 60% when comparing the benign with the malignant masses (See Table 61 for all threshold values).

Texture feature	Group pair	Threshold value	Sensitivity (%)	Specificity (%)
GLCM	Benign-malignant	279.5	75%	60%
	Malignant-cyst	177.5	100%	90%
	Benign-cyst	108	89%	78%
Wavelet	Benign-malignant	18284	50%	46%
	Benign-cyst	10982	75%	78%
	Malignant-cyst	11146	100%	78%

Table 61: Sensitivity and Specificity associated with GLCM and Wavelet features in premenopausal women

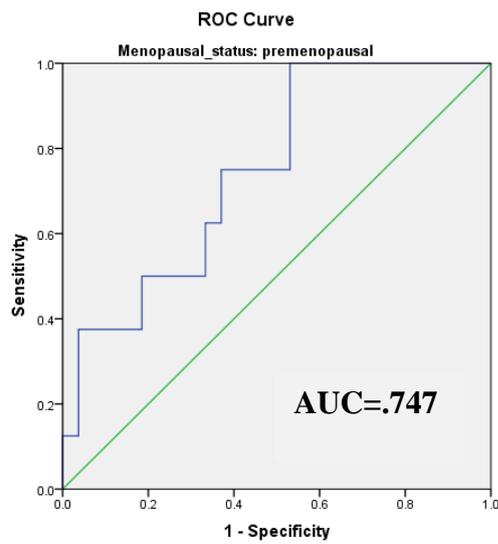


Figure 58: ROC curve for the GLCM feature for the difference between malignant and benign masses in the premenopausal group.

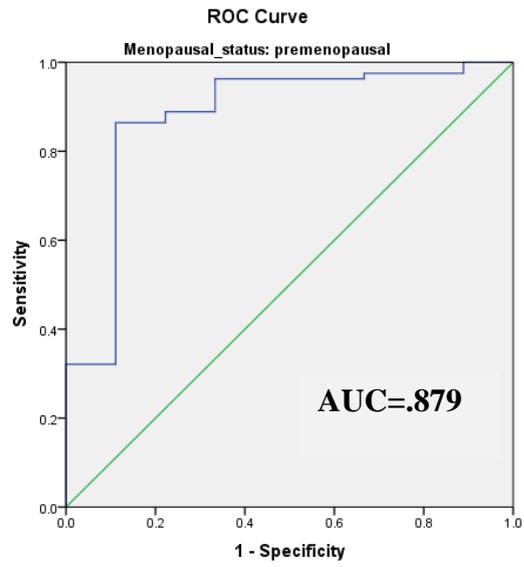


Figure 59: ROC curve for the GLCM feature for the difference between benign masses and cysts in the premenopausal group.

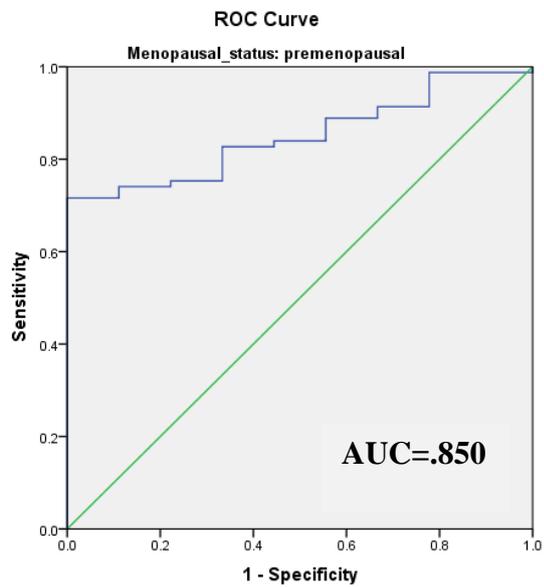


Figure 60: ROC curve for the wavelet feature for the difference between benign masses and cysts in the premenopausal group.

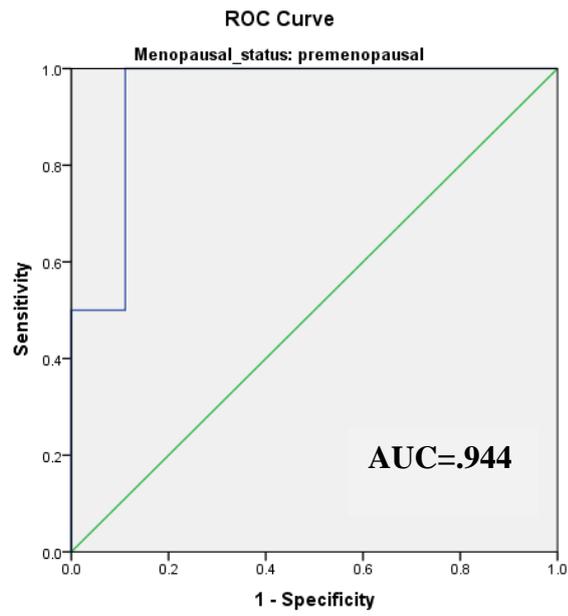


Figure 61: ROC curve for the GLCM feature for the difference between malignant masses and cysts in the premenopausal group.

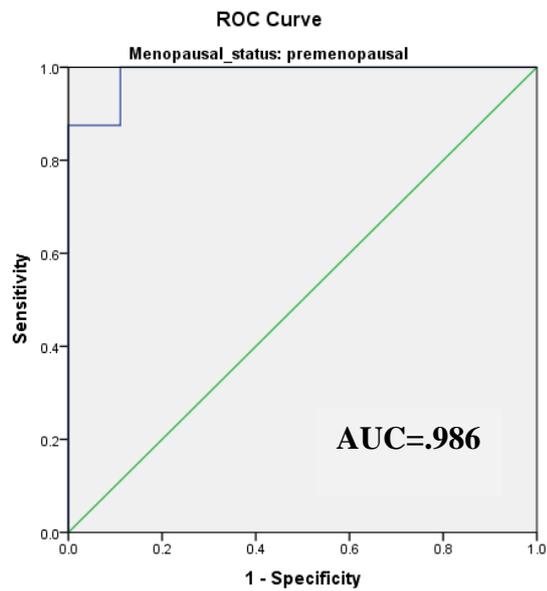


Figure 62: ROC curve for the wavelet feature for the difference between malignant masses and cysts in the premenopausal group.

4.3.6.1. PMI

When applying PMI scores in the premenopausal group, fifty-six masses were found. A contingency table was created and showed that eight masses were found to be malignant by histology. Six of them had a high or intermediate score in the PMI, which were the true positive results. Forty-eight masses were diagnosed as benign by histology and thirty-seven of them were classed as low risk on PMI, which were the true negative results. A summary of the results is showed in Table 62.

Type of the test	Malignant by histology	Benign by histology
High risk and intermediate risk	(TP) 6	(FP) 11
Low risk	(FN) 2	(TN) 37
Total	8	48

Table 62: 2x2 contingency table for PMI score in premenopausal women.

4.3.6.2. RMI

Here we applied RMI to the premenopausal group. Forty-nine masses were used for the analysis. Seven masses were diagnosed as malignant by histology; however, only one mass was found to have a high score in RMI, which was the true positive result. On the other hand, out of forty-two benign masses according to histology, forty had a low risk on the RMI score, which were the true negative results. Table 63 gives a summary of the results.

Type of the test	Malignant by histology	Benign by histology
RMI >250	(TP)	(FP)
	1	2
<250	(FN)	(TN)
	6	40
Total	7	42

Table 63: 2x2 contingency table for RMI score in premenopausal group

4.3.6.3. ADNEX

When applying the ADNEX scoring system to the premenopausal group, thirty-nine masses were analysed. A total of eight masses were found to be malignant by histology; four of them were found to have a high risk score by ADNEX, which were the true positive results. In the benign group, thirty-one masses were seen by histology. Twenty-five of them had a low risk score in the ADNEX and were the true negative results. Please see Table 64 for a summary of these results.

Type of the test	Malignant by histology	Benign by histology
High risk >50%	(TP)	(FP)
	4	6
Low risk <50%	(FN)	(TN)
	4	25
Total	8	31

Table 64: 2x2 contingency table for the ADNEX score in the premenopausal group

4.3.6.4. Sensitivity and specificity

The following table (Table 65) presents a summary of the diagnostic performance of the texture feature GLCM and the most commonly used scoring systems when analysing the premenopausal group. GLCM and PMI had the highest sensitivity (75%) compared to ADNEX and the wavelet feature (50%), while RMI had the lowest sensitivity (14%). However, RMI had the highest specificity (95%), followed by ADNEX (80%) and then GLCM (60%), and wavelet had the lowest specificity (46%). As can be seen, the sensitivity of wavelet, RMI, PMI and ADNEX (60%, 32%, 90% and 62% respectively in the total population) dropped when analysing the premenopausal group specifically, while the performance improved in the GLCM (sensitivity was 72% in total population). In the specificity performance, RMI improved from 87% to 95%, and PMI improved from 51% to 77%, while wavelet performance decreased from 60% to 48%. No change was seen in either GLCM or ADNEX.

Correlation	Sensitivity	Specificity	PPV	NPV	Accuracy
GLCM	75% (95% CI: 42-100%)	60%	15%	96%	60%
Wavelet	50% (95% CI: 1-99%)	46%	1%	90%	46%
RMI	14% (95% CI: too wide)	95%	33%	87%	84%
PMI	75% (95% CI: 42-100%)	77%	35%	95%	77%
ADNEX	50% (95% CI: 1-99%)	80%	40%	86%	74%

Table 65: Summary of the sensitivity, specificity, PPV, NPV and accuracy of the texture analysis feature GLCM compared to RMI, PMI and ADNEX in the premenopausal group

4.3.7. Postmenopausal group

Seventy-one women were found in the postmenopausal group, with thirty-five benign, twenty-one malignant and fifteen simple cysts. Similar to the total population analysis, group of pairs were compared to each other to test for significance. Significant differences were seen between the benign and the cyst group pair and between the malignant and the cyst group pair, but no significant difference was found between malignant and benign groups ($p=.110$). All group pairs in the wavelet feature showed a significant difference, unlike the results from the premenopausal group, where the difference between malignant and benign groups was not significant. Please see Table 66 for a summary of the results.

For those group pairs that demonstrate a statistically significance difference ($p<0.05$), ROC curves were created to determine the ability of the GLCM feature to discriminate between cysts and benign masses, and between cysts and malignant masses. For the wavelet feature, all three group pairs were analysed using a ROC curve. Please refer to Figures 64 to 68.

Area under the curve (AUC) results were somewhat similar to those for the premenopausal group. Excellent discriminatory ability was seen when comparing the cysts with the benign and the malignant masses when using the GLCM feature as well as between the cysts and the benign masses when using the wavelet feature. Good discriminatory ability between cysts and malignant masses was seen when using the wavelet feature. When comparing the benign with the malignant masses, the wavelet

feature showed poor discriminatory ability (AUC=.683) in the postmenopausal group.

A summary of the results is shown in Table 67.

Correlation	Sensitivity	Specificity	PPV	NPV	Accuracy
GLCM	75% (95% CI: 42-100%)	60%	15%	96%	60%
Wavelet	50% (95% CI: 1-99%)	46%	1%	90%	46%
RMI	14% (95% CI: too wide)	95%	33%	87%	84%
PMI	75% (95% CI: 42-100%)	77%	35%	95%	77%
ADNEX	50% (95% CI: 1-99%)	80%	40%	86%	74%

Table 66: Summary results of the significance test in postmenopausal women

Texture feature	Group pair	AUC	S.E.	Discriminatory ability
GLCM	Benign-malignant	.629	.085	Poor
	Benign-Cyst	.920	.043	Excellent
	Malignant-cyst	.943	.035	Excellent
Wavelet	Benign-malignant	.683	.079	Poor
	Benign-cyst	.768	.066	Fair
	Malignant-cyst	.863	.061	Good

Table 67: The area under the curve (AUC) and standard error (S.E.) associated with both GLCM and Wavelet features in postmenopausal women

Texture feature	Group pair	Threshold value	Sensitivity (%)	Specificity (%)
GLCM	Benign-malignant	267	71%	55%
	Benign-cyst	107.4	97%	70%
	Malignant-cyst	101.7	95%	73%
Wavelet	Benign-malignant	18283.9	62%	60%
	Benign-cyst	11104.9	71%	60%
	Malignant-cyst	10964.2	86%	60%

Table 68: Sensitivity and specificity associated with GLCM and Wavelet features in postmenopausal women

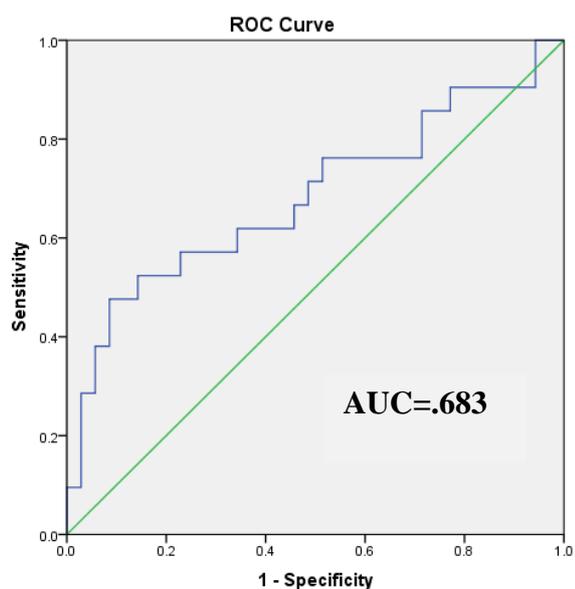


Figure 63: ROC curve for the wavelet feature for the difference between malignant and benign masses in the postmenopausal group

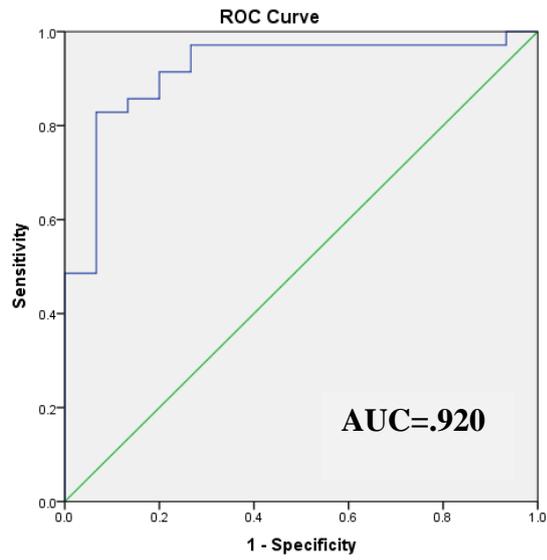


Figure 64: ROC curve for the GLCM feature for the difference between benign masses and cysts in the postmenopausal group

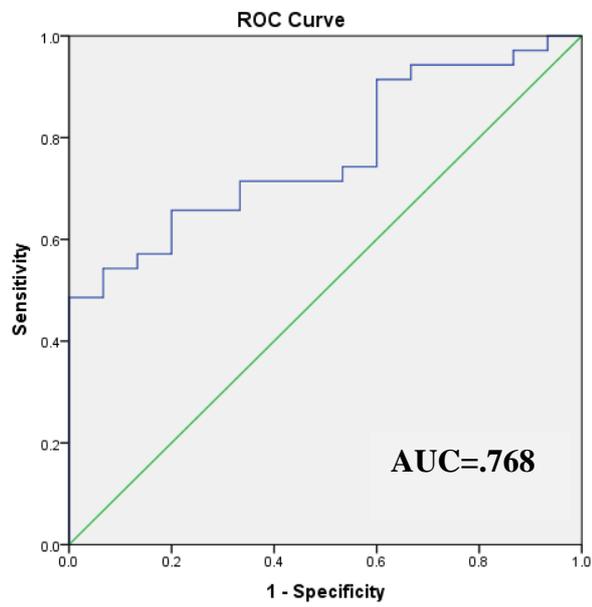


Figure 65: ROC curve for the wavelet feature for the difference between benign masses and cysts in the postmenopausal group

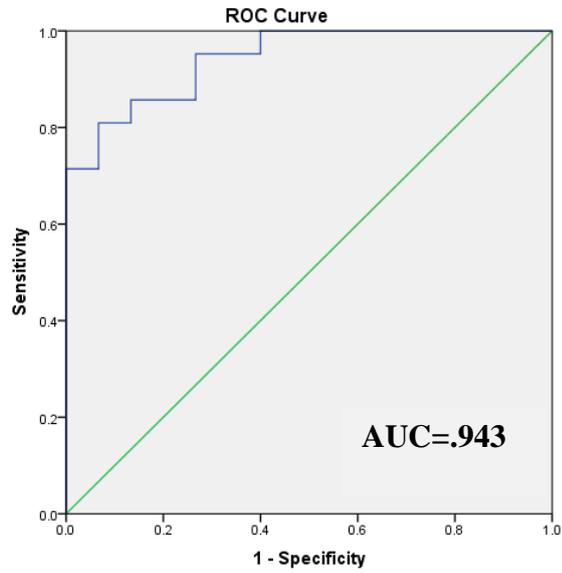


Figure 66: ROC curve for the GLCM feature for the difference between malignant masses and cysts in the postmenopausal group

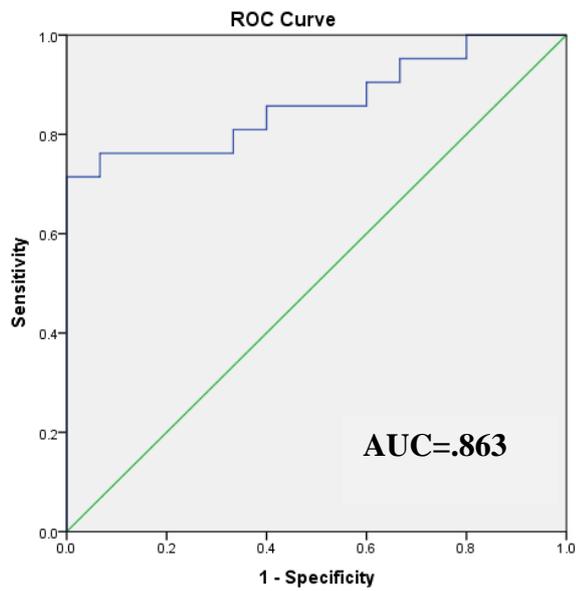


Figure 67: ROC curve of the wavelet feature for the difference between malignant masses and cysts in the postmenopausal group

4.3.7.1. PMI

In the postmenopausal group, thirty-nine masses were analysed using the PMI score. Nineteen of them were malignant and scored as high risk on the PMI: these were the true positive results. Six were benign and scored as low risk on the PMI: these were the true negative results. Table 69 is the contingency table created for the PMI in the postmenopausal group.

Type of the test	Malignant by histology	Benign by histology
High risk and intermediate risk	(TP) 19	(FP) 12
Low risk	(FN) 2	(TN) 6
Total	21	18

Table 69: 2x2 contingency table for PMI score in postmenopausal women.

4.3.7.2. RMI

In the postmenopausal group, twenty-nine masses were analysed using the RMI score. By histology, fifteen of them were found to be malignant. However, only six of these fifteen were classified as high risk using the RMI score (true positive results). The other fourteen masses were benign, and nine of them were true negative results: their RMI scores were <250, placing them in the low risk group. Table 70 summarises the results for the RMI scores in the postmenopausal group.

Type of the test	Malignant by histology	Benign by histology
High risk and intermediate risk	(TP) 19	(FP) 12
Low risk	(FN) 2	(TN) 6
Total	21	18

Table 70: 2x2 contingency table for RMI score in postmenopausal group

4.3.7.3. ADNEX

In the postmenopausal group, thirty-nine masses were analysed using the ADNEX scoring system. Twenty-one of them were found to be malignant by histology, and fourteen were high risk when using ADNEX (true positive results). The other seventeen were found to be benign by histology. Thirteen of them had low risk according to ADNEX, representing true negative results. Please refer to Table 71 for the contingency table of the ADNEX score in the postmenopausal group.

Type of the test	Malignant by histology	Benign by histology
High risk >50%	(TP) 14	(FP) 4
Low risk <50%	(FN) 7	(TN) 13
Total	21	17

Table 71: 2x2 contingency table for the ADNEX scores in the postmenopausal group

4.3.7.4. Sensitivity and Specificity

In the postmenopausal group, the PMI showed the best sensitivity (90%), followed by GLCM (71%) when compared to our texture analysis using the wavelet technique (48%) and to the other widely used scoring system, RMI (40%) and the new scoring system, ADNEX (66%). In general, all these methods improved in sensitivity performance when compared to the premenopausal group (75% for the PMI, 14% for RMI and 50% for ADNEX: please refer to Table 65 for the premenopausal group). Sensitivity decreased slightly for GLCM: from 75% to 71%. The specificity of the PMI was the lowest (33%). The highest specificity was seen in the ADNEX score (76%), followed by the wavelet (72%), and then the GLCM (55%). When compared to the premenopausal group, a noticeable decrease in specificity was seen: for example, in the PMI, specificity decreased from 77% to 33%; in the GLCM, it decreased from 60% to 55%; in the RMI, from 95% to 64%; and for the ADNEX, from 80% to 76%. A summary of these results is found in Table 72 for the postmenopausal group.

Correlation	Sensitivity	Specificity	PPV	NPV	Accuracy
GLCM	71% (95% CI: 49-93%)	55%	47%	75%	59%
Wavelet	48% (95% CI: too wide)	72%	62%	60%	60%
RMI	40% (95%CI: 0-79%)	64%	55%	50%	52%
PMI	90% (95% CI: 79-100%)	33%	61%	75%	64%
ADNEX	66% (95% CI: 29-100%)	76%	77%	65%	71%

Table 72: Summary of the sensitivity, specificity, PPV, NPV and accuracy of the Texture analysis feature wavelet compared to RMI, PMI and ADNEX in the postmenopausal group

In General, the ADNEX score was calculated for all masses, whether CA125 results were available or not. However, as mentioned by (Van Calster et al., 2014) in the study in which they developed the ADNEX model to differentiate between the different types of adnexal masses, CA125 was one of the strongest predictors, and they explained that deriving this model without the CA125 would decrease the discriminatory ability of the ADNEX. Therefore, it was decided to re-calculate the ADNEX score only for patients for whom CA125 results were available so that the difference could be appreciated.

It was found that of the eighty-one eligible masses, sixteen had missing CA125 results, divided equally between benign and malignant masses. Moreover, when dividing the population by menopausal status, six were in the premenopausal group and ten in the postmenopausal group. The ADNEX score was applied to sixty-five masses with available CA125 and resulted in 56% sensitivity, 81% specificity, 61% PPV, 77% NPV and 72% accuracy. A summary of the diagnostic performance of the ADNEX model for the total population as well as for pre- and postmenopausal groups is provided in Table 73. Surprisingly, the sensitivity of the ADNEX model decreased slightly when using only masses with available CA125: for example, in the total population, the ADNEX sensitivity decreased from 62% to 56%, and in the premenopausal group, from 50% to 40%. Lastly in the postmenopausal group, it decreased from 66% to 62%. However, the specificity, PPV, NPV and accuracy were similar to the ADNEX of all masses.

ADNEX with CA125 available	Sensitivity	Specificity	PPV	NPV	Accuracy
Total population	56% (95% CI 29-83%)	81%	61%	77%	72%
Premenopausal group	40% (95% CI 1-95%)	81%	38%	84%	74%
Postmenopausal group	62% (95% CI 32-92%)	80%	77%	66%	70%

Table 73: Summary of the sensitivity, specificity, PPV, NPV and accuracy of the ADNEX model with only CA125 available in all three groups of population.

4.3.8. Combination of the two texture analysis features

In order to improve the diagnostic performance, the two texture analysis features were combined, as confirmed by Singh and Singh (2002). Here, the same threshold values were used together to assess the diagnostic performance of the GLCM and the wavelet combined. Therefore, a threshold value of 245 was used for the GLCM and 17191 for the wavelet simultaneously to indicate the risk of malignancy. This was applied to 145 masses, of which 29 were malignant by histology and 116 were benign. Seventeen of the 29 malignant masses had values higher than 245 for GLCM and 17191 for the wavelet feature simultaneously: in other words, these were the true positive results, while 72 of the 116 benign masses were considered benign by texture analysis features combined, and represented true negative results. Please refer to Table 74 for a summary of these results.

When dividing the total population into pre- and postmenopausal groups, eighty-nine masses were used for the analysis of both textural features together in the

premenopausal group, of which eight were malignant and eighty-one were benign. Four of the eight masses that were malignant by histology were malignant by texture analysis as well: therefore, these were the true positive results. Meanwhile, forty-eight masses that were benign by histology were considered benign by texture analysis and subsequently were regarded as true positive results. Please refer to Table 75 for a summary of these results.

In the postmenopausal group, fifty-six masses were used for the analysis, of which twenty-one were malignant and thirty-five were benign. Thirteen of the twenty-one malignant masses were >245 in the GLCM feature and >17191 in the wavelet and therefore represent true positive results. Of the thirty-five benign masses, twenty-four were considered benign by texture analysis of both features together and represent true negative results. A summary of the data is shown in Table 76, below.

Type of the test	Malignant by histology	Benign by histology
GLCM >245 and wavelet	(TP)	(FP)
>17191	17	44
GLCM <245 and wavelet	(FN)	(TN)
<17191	12	72
Total	29	116

Table 74: 2x2 contingency table for the combined two texture analysis features.

Type of the test	Malignant by histology	Benign by histology
GLCM >279.5 and wavelet	(TP)	(FP)
>18284	4	33
GLCM <279.5 and wavelet	(FN)	(TN)
<18284	4	48
Total	8	81

Table 75: 2x2 contingency table for the combined two texture analysis features in
Premenopausal women

Type of the test	Malignant by histology	Benign by histology
GLCM >279.5 and wavelet	(TP)	(FP)
>18284	13	11
GLCM <279.5 and wavelet	(FN)	(TN)
<18284	8	24
Total	21	35

Table 76: 2x2 contingency table for the combined two texture analysis features in postmenopausal women.

4.3.4.1. Sensitivity and Specificity

In the total population, the sensitivity was 58% when combining the two texture analysis features together and slightly higher in the premenopausal group (61%), while the postmenopausal group yielded the lowest sensitivity (50%). When comparing these findings to the GLCM and wavelet analysis separately, it became

apparent that sensitivity decreased in all three groups (total population, pre- and postmenopausal). In addition, the specificities were more or less similar to the analysis of the features separately, ranging from 59% to 69%. Please refer to Table 77 for a summary of these results.

Correlation	Sensitivity	Specificity	PPV	NPV	Accuracy
Combined texture features in total population					
	58% (95% CI: 36-80%)	62%	27%	86%	61%
In premenopausal	50% (95% CI: 1-99%)	59%	11%	92%	58%
In postmenopausal	61% (95%CI: 35-87%)	69%	54%	75%	66%

Table 77: Summary of the sensitivity, specificity, PPV, NPV and accuracy of the texture analysis features (GLCM and wavelet) combined.

4.3.9. Logistic regression

Further analysis was performed on the data using logistic regression to explore the relationship of the variables to the outcome (histology results or follow-up). Testing for correlation between the variables was carried out to observe which variables were collinear so that they would not be used simultaneously in the equation. It was found that GLCM and wavelet were collinear, as were age and menopausal status. The final equation included menopausal status, wavelet and the ratio between wavelet and GLCM. This model was showed to have good predictive capacity when tested using the Hosmer and Lemeshow test ($p=.502$). Table 78 illustrate the results of the model. Moreover, ROC analysis was performed to determine the ability of this model to

discriminate between benign and malignant masses and gave an AUC of 0.81, which has a good discriminatory ability. Please refer to Figure 69 for the ROC.

Variable	Coefficient (β)	S.E.	P value	Odds Ratio	95% CI
Menopausal status	-1.834	.506	<0.05	.160	(.059 - .431)
Wavelet	.000	.000	<0.05	1.000	(1.000- 1.000)
Ratio	-.016	.007	0.032	.984	(.970 - .999)

Table 78: Logistic regression model result

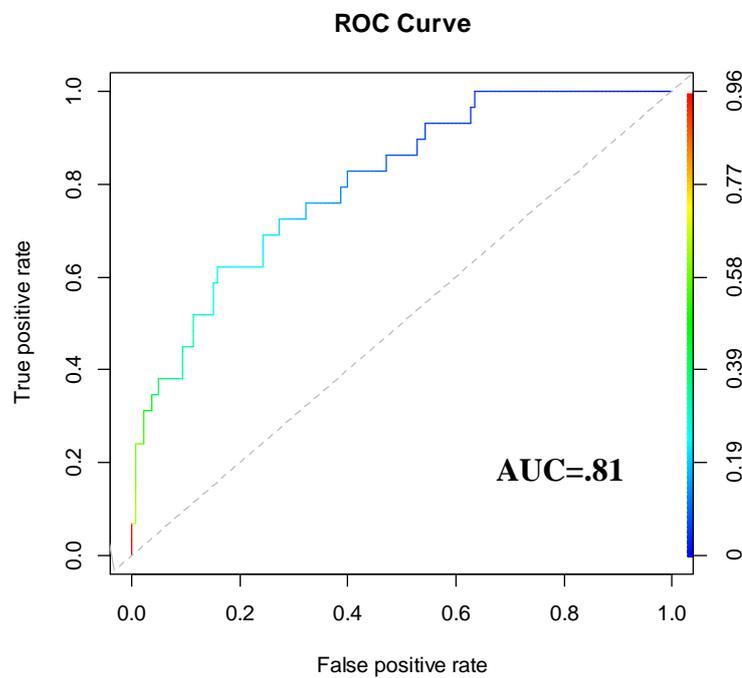


Figure 68: ROC for logistic regression model.

4.4. Summary

(A) GLCM

Although initially the GLCM showed a significant difference when comparing the benign and the malignant masses, the ROC curve showed a poor discriminatory ability, with $AUC=0.668$. Moreover, when dividing the study population into two groups (premenopausal and postmenopausal), the significant difference between the benign and the malignant masses was still seen in the premenopausal group using the GLCM feature for texture analysis. Here the ROC curve showed a fair ability to discriminate the benign from the malignant masses ($AUC=0.747$), while in the postmenopausal group, no significant difference was found between the benign and the malignant masses using GLCM.

GLCM showed a significant difference when comparing normal ovarian tissue with malignant tissue, with $p < 0.05$.

When dividing the benign masses into subgroups such as teratomas, endometriomas and fibroids, it was found that GLCM could significantly differentiate benign suspicious masses from teratomas and endometriomas, but it was statistically insignificant when comparing them to fibroids. This could be explained by the fact that most of the benign suspicious masses were fibromas or adenofibromas, which closely resemble fibroids in texture. On the other hand, when comparing these benign sub-groups with the malignant masses in order to establish whether GLCM could be helpful in this matter, a significant difference was seen between the malignant masses and both endometriomas and fibroids, but it could not differentiate benign masses

from teratomas, which could also be explained by the fact that they have a very similar texture to malignant masses.

(B) Wavelet

Similar to GLCM, the wavelet function showed a statistically significant difference between the three mass types: benign, malignant and cysts. However, when generating a ROC curve, the wavelet function showed a poor ability to discriminate benign from malignant masses, with $AUC=.630$. Moreover, when dividing the study population into premenopausal and postmenopausal groups, unlike the GLCM, the wavelet function did not show a significant difference between the benign and the malignant masses. In the postmenopausal group, the wavelet function showed a significance difference between the benign and the malignant masses. However, when generating a ROC curve, it revealed a poor ability to discriminate between benign and malignant masses, with $AUC=.683$

When looking at the difference between the benign subgroups, the wavelet function showed a significant difference between benign suspicious masses and teratomas, and between malignant masses and endometriomas. Unexpectedly, it could not differentiate between benign masses and endometriomas, although they have totally different texture features and can be easily distinguished visually. Similar to GLCM, the wavelet function could not differentiate benign from fibroid masses, as was expected.

However, when differentiating between normal ovarian tissue and malignant tissue, the wavelet function showed a significant difference, with $p=.004$.

4.5. Discussion

Ovarian cancer remains the leading cause of death among the gynaecological malignancies. To date, the nature of the mass – whether malignant or non-malignant – has to be confirmed by histology results, which means having a surgical procedure. Ultrasound imaging, on the other hand, is a non-invasive method for the diagnosis of ovarian cancer. Improvements in the diagnostic accuracy of the ultrasound will reduce the number of unnecessary surgeries.

The aim of this study was to determine the ability of GLCM and wavelet features in characterising ovarian tissue, and in particular, to investigate the diagnostic ability of these features in discriminating cysts and benign and malignant masses.

GLCM showed an excellent discriminatory ability to distinguish between malignant masses and cysts ($AUC=.994$) and between benign masses and cysts ($AUC=.895$), whilst for the wavelet feature it demonstrated a good ability of discrimination ($AUC=.894$ and $AUC=.814$ respectively). These findings are comparable to the results reported by (Hamid et al., 2011).

In their study involving image texture analysis of transvaginal ultrasound in monitoring ovarian cancer, Hamid et al. (2011) reported that the area under the ROC curve was .973 when comparing malignant masses with cysts and .889 when comparing normal tissue with cysts, and concluded that the GLCM feature can potentially be used to help in diagnosing ovarian cancer. Moreover, Hamid et al.

(2011) used 737 as a threshold value in GLCM, which led to an estimated sensitivity of 91% and specificity of 78%, while in our study, 86 was used as the threshold value, with an estimated sensitivity of 97% and specificity of 62%. The difference in the threshold values is obvious and led to an increase in our sensitivity to differentiate malignant masses from cysts, with a slight decrease in specificity.

For the wavelet feature, Hamid and colleagues used 4,362 as the threshold value with both sensitivity and specificity set at 90%, while in the current study 10,484 was used as the threshold value, which led to similar sensitivity (90%) but decreased specificity (59%).

In our study, we observed tissue characterisation using texture analysis by comparing normal ovarian tissue with solid areas of malignant masses. A significant difference was noticed when using both GLCM and wavelet features. These results are comparable to those of Hamid et al. (2011). Furthermore, it is worth mentioning that in our study, the technique for drawing the ROI on the ultrasound image was adopted from Hamid et al. (2011). He excluded anechoic areas from the ROI of complex masses after comparing this approach with including the whole mass in the ROI and found that there was a significant difference and that accuracy was improved when excluding the anechoic areas.

Our results showed no significant difference between malignant and teratoma when using both GLCM and wavelet. This could be explained by the fact that teratoma usually contains bone and teeth (solid areas) that could resemble in texture to

malignant solid masses. However a significant difference was seen between malignant and fibroids using GLCM which could be useful clinically in the future.

A diagnostic method should be positive for those with the disease (sensitivity) and negative for those without the disease (specificity). Our data analysis showed that the sensitivity of the GLCM feature using a threshold value of 245 as indicative of malignancy was 72%, which means that the ability of this texture feature to detect correctly those who have ovarian malignancy (true positive) is 72% and the remaining 28% are false positive results. On review, the majority of the falsely categorised tumours were cystadenomas and fibromas, which are known to be difficult to characterise on ultrasound (Valentin, 2000, Valentin, 2004).

Our sensitivity of GLCM is considered low: the reason could be the fact that the appearance of many benign lesions overlaps with that of malignant diseases (Varras, 2004).

These sensitivity results are lower than those been published in earlier studies: for example, sensitivity was 92% in (Xian, 2010) study, which applied the GLCM texture feature to identify malignant and benign liver tumours on ultrasound images. This difference in sensitivity could be explained by the fact that texture analysis is more appropriate for the characterisation of regions exhibiting homogeneity in their structure, as discussed by Diamond et al. (2004) in their study of liver tissue, compared to the less homogenous ovarian masses in our study.

Similarly, a higher sensitivity of 93%, specificity of 86% and AUC of .956 were achieved in a study published in 2007 by Michail and colleagues, in which they

studied the texture analysis of pre- and postmenopausal endometrial tissue via transvaginal ultrasound (Michail et al., 2007).

For the wavelet feature, we used a threshold value of 17191 as indicative of malignancy, which led to 60% sensitivity: therefore, the ability of this texture feature to detect correctly ovarian masses (true positive) is 60%, and the remaining 40% are false positive results.

Moreover, the ability of the wavelet feature to rule out those who do not have ovarian malignancy was 60%: this specificity means that 60% of the non-diseased women gave a true negative result.

Predictive value is an important measure for diagnostic tests. Predictive values measure whether or not the individual actually has the disease. Positive predictive value (PPV) tells a postmenopausal woman how likely is it that she has an ovarian malignancy if she has an analysis of >279 in the GLCM feature. In our analysis, PPV of GLCM in the postmenopausal group was 47%, which indicates that the probability that a woman with GLCM value of >279 actually has an ovarian malignancy is 47%. This probability limits unnecessary surgery to its half number without using the technique.

Our GLCM PPV in the postmenopausal group (47%) is lower than the PPV that has been previously reported: 91.8% (Xian, 2010) and 99.7% (Acharya et al., 2013). Again, this could be as a result of thier studies are applied for other organs. On the other hand, a study conducted to evaluate the performance of pelvic mass index PMI (not a texture analysis feature) reported a lower PPV of 31.4% (Sinha A et al., 2015).

These results could be also explained by the low number of women diagnosed with ovarian cancer (29.6%) among our study postmenopausal group (postmenopausal women who had complex masses by ultrasound and experienced surgical intervention), which is lower than previous studies, which documented a prevalence of 34% (Xian, 2010) but higher than the PMI study, which documented a prevalence of 16.7% (Sinha A et al., 2015). Nevertheless, the more sensitive the feature, the less likely it is that women with a GLCM of less than 245 will have ovarian malignancy, and so the greater the negative predictive value.

Our analysis showed a NPV of 90% in the total population, which indicates that the probability that a women with a GLCM value of less than 245 does not have an ovarian malignancy is 90%. This finding is higher than the 37.5% described by Zimmer et al. (2003), but is close to the previous PMI study NPV of 96% (Sinha A et al., 2015) and lower than the 98% documented by (Xian, 2010).

In a study by Vidya et al. (2015) using computer aided diagnosis in ultrasound to diagnose myocardial infarction (MI) using three different types of texture analysis, including GLCM, it was found that GLCM achieved an accuracy of 85%, sensitivity of 90% and specificity of 81%. Moreover, another recent study on breast morphologic features that focused on developing a computer-aided diagnosis (CAD) system based on texture features for distinguishing between benign and malignant breast cancers concluded that texture analysis can improve the ability to discriminate between benign and malignant breast lesions, with sensitivity of 96% (Moon et al., 2015). These studies yielded higher sensitivity and accuracy than our results, which showed a sensitivity of 72%, specificity of 60% and an accuracy of 60% for GLCM. The

reason for this difference could be explained by the larger samples used in their studies – 400 MI and 400 normal, compared to 116 benign and 29 malignant – which gave better results. However, collecting a similar number of images was not possible in our study due to the short period of recruitment approved by the ethics committee for this study (18 months) and the low prevalence of ovarian cancer in the population, as explained by Sinha A et al. (2015), who collected only 56 cancers compared to 414 benign masses over a period of seven years, which makes it nearly impossible to reach this large number using only single centre data.

Our study showed that GLCM had better performance than both the RMI and the ADNEX model when differentiating benign from malignant masses, even when applying them to pre- and postmenopausal groups separately. Although RMI should have improved performance when applied to the postmenopausal group, it still showed the lowest sensitivity (40%) among all the scoring systems applied. Our results showed a much lower performance of RMI (ranging from 40-14%) when compared with previous studies that validate the RMI. For example, the most recent systematic review and meta-analysis to investigate the diagnostic ability of several scoring systems calculated the pooled sensitivity and specificity of RMI at 72% and 92% respectively (Kaijser et al., 2014).

Amongst our patient cohort, the RMI had the highest specificity (87%); however, as reported by Myers and colleagues, studies on tests with high sensitivity usually report a lower specificity and vice versa (Myers ER. et al., 2006).

Similarly, our results demonstrate a lower sensitivity (62%) of the ADNEX model when compared to the original IOTA study, which concluded 94% sensitivity with almost equal specificity to our results, at 77% (Van Calster et al., 2014).

The difference in sensitivity could be explained by Mol and colleagues in their study, in which they externally validated the prognostic models that were used to distinguish benign from malignant adnexal masses, concluding that the diagnostic performance of these models is not as good as that reported in the original publication (Mol et al., 2001).

Another obvious factor that contributes to this variety in performance could be the much smaller sample that was used in our analysis (81 masses, of which 29 were malignant and 52 were benign, with 35.8% prevalence), compared to the ADNEX study's huge sample population of 5909 women, with 3980 benign and 1929 malignant masses, with prevalence between 22%-66% in the study oncology centres (Van Calster et al., 2014).

When applying PMI, our results agree with those of Sinha A et al. (2015) that PMI yields the highest sensitivity in our patient cohort in all three groups: 90%, 75% and 90% for the total population, pre- and postmenopausal groups respectively. On the other hand, our results showed a lower specificity of PMI (ranging between 51-33%) than in their study (60%).

Although results showed higher sensitivity of PMI (90%) compared to GLCM (75%), it has to be said that GLCM has the advantage of being objective, compared to the subjectivity of the PMI score, which is highly operator dependent. In addition, both

have the advantage over the RMI and the ADNEX in the fact that they are not dependent on the CA125.

According to (Materka, 2004, Tesař et al., 2008), the GLCM feature is commonly used in 2D texture analysis of medical images. The results from this study demonstrated that in general, the GLCM has a better characterisation ability compared to the wavelet feature. This is in agreement with several previous studies: for example, Hamid et al. (2011), in their pilot study to investigate the use of texture analysis features in monitoring ovarian masses, concluded that GLCM had a higher performance than other texture features. This is in accordance with the statement by Tuceryan and Jain (1998) that GLCM generally outperforms other features. Likewise, in a study that focused on breast lesions, Garra et al. (1993) reported that GLCM is the most useful feature and stated that GLCM has long been a powerful tool for texture analysis.

It was recommended in previous studies that combining texture features yields better performance compared to using features from a single category (Singh and Singh, 2002, Michail et al., 2007). In our study, we applied this theory by combining GLCM with the wavelet feature to analyse the masses. Unexpectedly, this method did not yield better sensitivity nor specificity compared to using GLCM alone: in fact, it showed decreased performance in the total population as well as in both pre- and postmenopausal groups. This could be explained by the fact that the GLCM feature did not show a significant difference initially in differentiating between benign and malignant masses in the postmenopausal group, and therefore, when adding it to the wavelet feature analysis, the diagnostic performance of the combined features

decreased. The same concept is applied in the premenopausal group, where the performance of the test decreased when combining the GLCM with the wavelet feature; however, here the wavelet was the feature that gave the non-significant results initially when differentiating benign from malignant masses in the premenopausal group.

Lastly, we explored the possibility of combining these two features in a different statistical way. This was applied by using them in a logistic regression model, which revealed that using the wavelet feature with the ratio of wavelet to GLCM along with the menopausal status as variables of the model gave a better discriminatory ability of AUC=0.8 and had a good predictive capacity when tested using the Hosmer and Lemeshow test ($p=.502$).

5. General discussion, Conclusion, Limitations and future studies

5.1 General discussion and Conclusion

Ovarian cancer accounts for about three percent of all cancers in women and is the fifth leading cause of cancer-related death among women in the United States. In 2014, it is estimated that nearly 22,000 women will be diagnosed with ovarian cancer in the United States, and approximately 14,000 will die of the disease. (National Cancer Institute, 2014)

In Europe, around 65,600 new cases of ovarian cancer were estimated to have been diagnosed in 2012. The UK incidence rate is the ninth highest in Europe. Worldwide, nearly 239,000 women were estimated to have been diagnosed with ovarian cancer in 2012, with incidence rates varying across the world (Cancer Research, 2015).

Ovarian cancer causes more deaths than any other cancer of the female reproductive system. This high mortality rate is because of the absence of early symptoms and a lack of effective screening tests. Consequently, ovarian cancer is often diagnosed at an advanced stage, after the disease has spread outside the ovary (National Cancer Institute, 2014).

Among the various types of imaging modalities, ultrasound is considered the main imaging procedure for scanning ovarian masses. However, the main issue with ultrasound is that it is operator-dependent, and thus the accuracy and reproducibility of the diagnosis vary according to the experience of the operator.

In order to reduce the operator dependency, texture analysis, which is able to quantitatively characterise tissue through texture content, has been considered in this study. In medical imaging, texture analysis has been applied after proving to be valuable in characterising tissue such as liver, breast, prostate, carotid plaques and many others.

As mentioned in Chapter 1, the aim of this study was to determine the ability of texture analysis features, namely GLCM and wavelet, as well as the new ASQ feature, in characterising ovarian tissue. To the best of the author's knowledge, this study is the first to attempt to apply ASQ to ultrasound images of ovarian masses.

5.1.1. ASQ

Before applying ASQ to images of pelvic masses, it had to be tested in several aspects, since it has not previously been used in pelvic masses. Therefore, we tested the variability and reliability of ASQ on ultrasound images caused by random variation during image acquisition. It was found that the ASQ feature demonstrated excellent repeatability for ASQ software, with all transducers showing less than 0.4% variance from the mean, which indicates that ASQ software is able to produce reliable ASQ output measures.

In this study, ASQ was applied to 45 pelvic masses on ultrasound images to investigate the ability of this method to distinguish between benign and malignant pelvic masses. The preliminary results showed no significant difference between benign and malignant masses using the ASQ technique. For this reason, recruiting more patients was pointless and the study was terminated. This failure to discriminate

the benign from the malignant masses using ASQ could be because most of the masses in both groups had similar fibrotic nature, such as cystadenofibromas, fibroids and adenocarcinomas. Another reason could be the restricted ROI drawing, which did not allow us to include the whole mass in the image: instead, the ROI was drawn as a fan shape.

In this study, two different techniques were adopted to draw the ROI. First, the whole mass was included in the ROI as much as possible. The second technique was to draw the ROI only on the solid area of the mass. Nevertheless, using different techniques did not contribute positively in distinguishing benign from malignant masses.

ASQ works well on liver images and has proven to be beneficial in diagnosing various liver diseases (Brosky, 2009, Toyoda et al., 2009, Hung, 2010, Wang et al., 2013, Ricci et al., 2013, Onodera, 2013). For this reason, we thought it might also work well on ovarian masses.

Unfortunately, our study showed negative results: the quantitative ASQ analysis of B-mode images demonstrated non-significant differences between benign and malignant tissue. This means that ASQ does not work on pelvic masses because both benign and malignant masses exhibit homogeneity and heterogeneity in the same way. This could be explained by the fact that liver is a smooth homogenous organ while ovary is in general normally heterogeneous due to the presence of follicles and the physiological changes that occur in it each cycle.

5.1.2. Texture analysis

To carry out a good study, larger sample sizes should give more reliable results with greater precision and power; however, they also cost more time and money. Therefore, it is important to perform a sample size calculation before conducting a study to ensure a sufficiently large sample size to be able to draw meaningful conclusions, without wasting resources on sampling more than what really needed (Select Statistical, 2015).

In this study, the calculation of sample size was based on the assumption that texture analysis may distinguish between benign and malignant masses, and after seeking statistical advice, it was concluded that with a 5% significance level and 80% power, a total of 200 women (100 participants in each arm) were required for the study to demonstrate a significant difference. Unfortunately, the total number of masses that were analysed in our study was 169, of which 140 were benign and 29 were malignant by histology results.

It was understood that the number of malignant masses was too small; however, the recruitment of such masses was difficult for several reasons. First, most women with malignancy tend to be more ill than others with benign lesions, and therefore, the majority refused to participate in the study and have an additional internal scan. Second, some malignant diseases were disseminated and were thus excluded from the study because we were unable to view them on ultrasound images and analyse them. Third, the probability of ovarian cancer is higher in older postmenopausal women

(>70 years old) which makes it difficult to approach these women and ask for their participation in the study.

Our study showed that the GLCM and wavelet texture analysis features are able to distinguish between benign and malignant masses with sensitivities of 72% and 60% respectively. The use of texture analysis in the diagnosis of ovarian tumours has some advantages: for example, the end results are more objective and reproducible compared to manual interpretation of ultrasound images, which can occasionally be affected by inter-observer variation. Another advantage is that since we use images acquired using the commonly available and affordable ultrasound modality, there is no additional cost for image acquisition. Moreover, the MaZda software that is used for texture analysis is available online free of charge and can be downloaded easily to any computer. The operator just has to run the software on the acquired B-mode ultrasound image and the software does all the processing and gives the output results after characterising the tissue. Hence, there is no need for trained experts to run the software.

As proposed by Huynen et al. (1994), texture analysis may improve diagnostic accuracy by providing more reproducible results and information that is difficult for humans to comprehend.

This study demonstrated that texture analysis methods, particularly GLCM and wavelet features, are potentially able to discriminate between normal and pathological ovarian tissue. It is worth noting that the threshold value should be obtained by performing ROC curve analysis on a large sample size that is represented by the group studied before a threshold value can be set for clinical studies.

Our results showed a significant difference between malignant and fibroid using GLCM feature which can be useful clinically in the future, such as in some cases of degenerated fibroid that are difficult to diagnose visually by the operator.

In our study, diagnostic performance of two texture analysis features namely GLCM and wavelet as well as the widely used scoring systems (RMI, PMI and ADNEX) were applied to the total sample population. The GLCM showed a higher sensitivity (72%) compared to two of the scoring systems applied (32% RMI and 62% ADNEX).

In addition, analysis was carried out in the premenopausal and postmenopausal groups separately. Interestingly, GLCM and wavelet diagnostic performance was better in premenopausal compared with postmenopausal women (75% vs. 71% in GLCM and 50% vs. 48% in wavelet), in contrast to the performance of RMI, PMI and ADNEX, which improved when applied to postmenopausal women (14% vs. 40% in RMI, 75% vs. 90% in PMI and 50% vs. 66% in ADNEX). This gives texture analysis an advantage over the other scoring systems when dealing with premenopausal women.

GLCM and wavelet can also be useful tools to distinguish between malignant and cystic masses and between benign and cystic masses, since our results showed an excellent discriminatory ability of GLCM (AUC=.994; AUC=.895) and good ability in the wavelet feature (AUC=.894; AUC=.814) respectively.

Another significant result was observed in our study when comparing normal ovarian tissue with solid areas of malignant masses when applying both GLCM and wavelet features, with $p < .05$ and $p = .004$ respectively.

Generally, our results demonstrate that GLCM has a better characterisation ability compared to the wavelet feature.

Moreover, combining the two features of texture analysis the GLCM and the wavelet feature did not seem to give a better diagnostic performance. In fact, a decrease in performance was noticed; however, combining these two features in a different statistical way using a logistic regression model gave a better discriminatory ability. The developed logistic regression model produced using wavelet, the ratio of wavelet to GLCM, and the menopausal status as variables in the model resulted in AUC=0.8 and a good predictive capacity when tested using the Hosmer and Lemeshow test ($p=.502$).

The performance of the texture analysis methods is determined by measuring how they articulate the relationship between image elements (AL-KADI, 2009). Since the way in which a texture analysis method calculates the image parameters was not the primary objective of this study, we did not investigate the underlying factors which cause the difference in the performance of the GLCM and wavelet features. Lerski (2006) explained that the whole relationship of texture parameters to tissue structure is a very complex issue which is not fully understood.

5.2. Study limitations

It is clear that in this study we analysed a relatively small sample of ovarian masses (29 malignant masses), as mentioned in the previous chapter. It is worth mentioning that at the beginning of the recruitment period (the whole first year, to be exact), the Gynaecology oncology clinic and the pelvic mass clinic from which we recruit participants were situated at different hospitals than the one where we scanned the patients using the Toshiba Aplio machine that has the ASQ feature. Thus, participants might not have been keen on taking part in a study for which they would have had to travel at their own expense to another hospital with payable parking to volunteer to have a scan for the study, which in turn resulted in a smaller number of participants than anticipated in this study.

This, in addition to the limited duration of recruitment time that was approved by the ethical committee (18 months) and the fact that data were collected from a single centre, all contributed to our inability to achieve the sample size that was desired.

Although texture analysis is a method applied to extract additional information from medical images that is difficult to apprehend by visual inspection, such analysis remains limited by the restricted resolution of the images.

In the ASQ technique, the major limitation was the restriction in drawing the ROI in the ASQ software. It only allows ROI to be drawn in a fan shape, which made it difficult to include the whole pelvic mass of different shapes without including other tissue or omitting part of the mass.

5.3. Future work

To assess the reproducibility of texture analysis features on ultrasound images of ovarian masses, a multicentre study should be carried out prospectively with a larger study sample to get more representative results and to confirm the clinical importance of this technique.

Additionally, it might be beneficial to explore the possibility of combining texture analysis with Doppler flow assessment to achieve higher diagnostic performance in discriminating ovarian tissue as it might be the advantage of PMI method is that it includes Doppler information.

Moreover, combining the GLCM and wavelet in other ways to achieve better results, such as the use of artificial neural networks, and combining GLCM/wavelet with scoring systems has the potential to improve the research, or using these features in a logistic regression model with a larger sample all could be used to improve this current work by applying it in future work.

Lastly, ovarian cancer is classified into four stages based on the FIGO system. Future work may involve the classification of malignant tissue using texture analysis techniques at different stages similar to what has been done using the ADNEX model except with the advantage of being an objective rather than a subjective assessment.

Presentations and Posters

- ASQ technology, PGR conference in Cardiff University 19th February 2014, poster.
- ASQ technology, Medical Physics and Clinical Engineering in Wales Summer Meeting, Cardiff, UK 27th June 2014, Presentation.
- The application of ASQ in the diagnosis of pelvic masses (preliminary study), BMUS conference in Manchester, UK 8th December 2014, Presentation. Published: RANA ALDAHAWI, NEIL PUGH. Proceedings of the British Medical Ultrasound Society 46th Annual Scientific Meeting 9–11 December 2014, *The application of ASQ*, Ultrasound May 2015 vol. 23 no. 2 NP7.
- ASQ technology, 8th Saudi Student Conference SSC in London, UK 30TH January 2015, poster.
- The use of texture analysis in diagnosing ovarian masses (Preliminary results), BSGI meeting in London 14th April 2015, presentation. Published: RANA ALDAHLAWI, NEIL PUGH: *The use of texture analysis in diagnosing ovarian masses (preliminary results)*. BJOG **British journal** of obstetrics and Gynaecology June 2015 volume 122, issue supplement S3 page 3.
- The Use of Texture Analysis in Diagnosing Ovarian Masses, BMUS conference in Cardiff, UK 8th December 2015, presentation.
- Texture analysis of transvaginal images in the diagnosis of ovarian cancer, EBCOG 14th European congress in Torino, Italy, 19-21 May 2016, Poster.

References

ACOG Committee Opinion, Number 247, December 2000. *International Journal of Gynaecology & Obstetrics*, 82, 241-245.

ACHARYA, U. R., SREE, S. V., SABA, L., MOLINARI, F., GUERRIERO, S. & SURI, J. S. 2013. Ovarian tumour characterization and classification using ultrasound - A new online paradigm. *Journal of Digital Imaging*, 26, 544-553.

AGARWAL, A., REIN, B. J. D., GUPTA, S., DADA, R., SAFI, J. & MICHENER, C. 2011. Potential markers for detection and monitoring of ovarian cancer. *Journal of Oncology*.

AGUIRRE, A., ARDESHIRPOUR, Y., SANDERS, M. M., BREWER, M. & ZHU, Q. 2011. Potential role of coregistered photoacoustic and ultrasound imaging in ovarian cancer detection and characterization. *Translational Oncology*, 4, 29-37.

AGUIRRE, A., GUO, P., GAMELIN, J., YAN, S., SANDERS, M. M., BREWER, M. & ZHU, Q. 2009. Coregistered three-dimensional ultrasound and photoacoustic imaging system for ovarian tissue characterization. *Journal of Biomedical Optics*, 14.

AKDENIZ, N., KUYUMCUOĞLU, U., KALE, A., ERDEMOĞLU, M. & CACA, F. 2009a. Risk of malignancy index for adnexal masses. *European Journal of Gynaecological Oncology*, 30, 178-180.

AKDENIZ, N., KUYUMCUOĞLU, U., KALE, A., ERDEMOĞLU, M. & CACA, F. 2009b. Risk of malignancy index for adnexal masses. *European Journal of Gynaecological Oncology*, 30, 178-180.

AKTÜRK, E., KARACA, R. E., ALANBAY, I., DEDE, M., KARASHAHIN, E., YENEN, M. C. & BAŞER, I. 2011. Comparison of four malignancy risk indices in the detection of malignant ovarian masses. *Journal of Gynaecologic Oncology*, 22, 177-182.

AL-KADI, O. S. 2009. Tumour Grading and Discrimination based on Class Assignment and Quantitative Texture Analysis Techniques. PhD in Engineering, University of Sussex.

ALACAM, B., YAZICI, B. & BILGUTAY, N. Year. Breast cancer detection based on ultrasound B-scan texture analysis and patient age information. In: *Bioengineering Conference, 2003 IEEE 29th Annual, Proceedings of, 22-23 March 2003* 2003. 98-99.

ALCÁZAR, J. L. 2006. Tumour angiogenesis assessed by three-dimensional power Doppler ultrasound in early, advanced and metastatic ovarian cancer: A preliminary study. *Ultrasound in Obstetrics and Gynaecology*, 28, 325-329.

ALCÁZAR, J. L., AUBÁ, M. & OLARTECOECHEA, B. 2012a. Three-dimensional ultrasound in gynaecological clinical practice. *Reports in Medical Imaging*, 5, 1-13.

ALCÁZAR, J. L. & CASTILLO, G. 2005. Comparison of 2-dimensional and 3-dimensional power-Doppler imaging in complex adnexal masses for the prediction of ovarian cancer. *American Journal of Obstetrics and Gynaecology*, 192, 807-812.

ALCÁZAR, J. L., ERRASTI, T., LAPARTE, C., JURADO, M. & LÓPEZ-GARCÍA, G. 2001. Assessment of a new logistic model in the preoperative evaluation of adnexal masses. *Journal of Ultrasound in Medicine*, 20, 841-848.

ALCÁZAR, J. L. & GUERRIERO, S. 2011. Grey-scale ultrasound versus CA-125 levels for predicting malignancy in adnexal masses. *International Journal of Gynaecology and Obstetrics*, 114, 290-291.

ALCÁZAR, J. L., GUERRIERO, S., LAPARTE, C., AJOSSA, S. & JURADO, M. 2011. Contribution of power Doppler blood flow mapping to grey-scale ultrasound for predicting malignancy of adnexal masses in symptomatic and asymptomatic women. *European Journal of Obstetrics & Gynaecology and Reproductive Biology*, 155, 99-105.

ALCÁZAR, J. L., GUERRIERO, S., PASCUAL, M. A., AJOSSA, S., OLARTECOECHEA, B. & HERETER, L. 2012b. Clinical and sonographic features of uncommon primary ovarian malignancies. *Journal of Clinical Ultrasound*, 40, 323-329.

ALCÁZAR, J. L., ITURRA, A., SEDDA, F., AUBÁ, M., AJOSSA, S., GUERRIERO, S. & JURADO, M. 2012c. Three-dimensional volume off-line analysis as compared to real-time ultrasound for assessing adnexal masses. *European Journal of Obstetrics Gynaecology and Reproductive Biology*, 161, 92-95.

ALCÁZAR, J. L. & JURADO, M. 2011. Three-dimensional ultrasound for assessing women with gynaecological cancer: A systematic review. *Gynaecologic Oncology*, 120, 340-346.

ALCÁZAR, J. L., MERCÉ, L. T., LAPARTE, C., JURADO, M. & LÓPEZ-GARCÍA, G. 2003. A new scoring system to differentiate benign from malignant adnexal masses. *American Journal of Obstetrics and Gynaecology*, 188, 685-692.

ALCÁZAR, J. L., MERCÉ, L. T. & MANERO, M. G. 2005. Three-dimensional power Doppler vascular sampling: A new method for predicting ovarian cancer in vascularized complex adnexal masses. *Journal of Ultrasound in Medicine*, 24, 689-696.

ALCÁZAR, J. L., OLARTECOECHEA, B., GUERRIERO, S. & JURADO, M. 2013a. Expectant management of adnexal masses in selected premenopausal women: A prospective observational study. *Ultrasound in Obstetrics and Gynaecology*, 41, 582-588.

ALCÁZAR, J. L., PASCUAL, M. Á., OLARTECOECHEA, B., GRAUPERA, B., AUBÁ, M., AJOSSA, S., HERETER, L., JULVE, R., GASTÓN, B., PEDDES, C., SEDDA, F., PIRAS, A., SABA, L. & GUERRIERO, S. 2013b. IOTA simple rules for

discriminating between benign and malignant adnexal masses: Prospective external validation. *Ultrasound in Obstetrics and Gynaecology*, 42, 467-471.

ALCAZAR, J. L., ROYO, P., JURADO, M., MÃNGUEZ, J. A., GARCÃA-MANERO, M., LAPARTE, C., GALVÃN, R. & LÃ³PEZ-GARCÃA, G. 2008. Triage for surgical management of ovarian tumours in asymptomatic women: Assessment of an ultrasound-based scoring system. *Ultrasound in Obstetrics and Gynaecology*, 32, 220-225.

ALETTI, G. D., GALLENBERG, M. M., CLIBY, W. A., JATOI, A. & HARTMANN, L. C. 2007. Current management strategies for ovarian cancer. *Mayo Clinic Proceedings*, 82, 751-770.

ALQAHTANI, M. A., COLEMAN, D. P., PUGH, N. D. & NOKES, L. D. M. Year. Tissue characterization: Influence of ultrasound setting on texture features in vivo. In, 2010. 63-66.

ALTMAN, D. 1995. *Practical statistics for medical research*, London, Chapman and Hall.

AMD. 2010. Prostate Histoscanning [Online]. Available: <http://www.histoscanning.com/products/prostate-histoscanningtm> / [Accessed 8/8 2012].

AMD. 2010. Ovarian Histoscanning [Online]. Available: <http://www.histoscanning.com/products/in-development/ovarian-histoscanningtm> / [Accessed 8-8 2012].

AMERICAN CANCER SOCIETY, W. 2015. Cancer facts and figures 2015 [Online]. Available: <http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2015> / [Accessed 10 February 2016].

AMEYE, L., VALENTIN, L., TESTA, A. C., VAN HOLSBEKE, C., DOMALI, E., VAN HUFFEL, S., VERGOTE, I., BOURNE, T. & TIMMERMAN, D. 2009. A scoring system to differentiate malignant from benign masses in specific ultrasound-based subgroups of adnexal tumours. *Ultrasound in Obstetrics and Gynaecology*, 33, 92-101.

AMONKAR, S. D., BERTENSHAW, G. P., CHEN, T. H., BERGSTROM, K. J., ZHAO, J., SESHIAH, P., YIP, P. & MANSFIELD, B. C. 2009. Development and preliminary evaluation of a multivariate index assay for ovarian cancer. *PLoS ONE*, 4.

AMOR, F., ALCÁZAR, J. L., VACCARO, H., LEÓN, M. & ITURRA, A. 2011. GI-RADS reporting system for ultrasound evaluation of adnexal masses in clinical practice: A prospective multicentre study. *Ultrasound in Obstetrics and Gynaecology*, 38, 450-455.

- ANDERSEN, E. S., KNUDSEN, A., RIX, P. & JOHANSEN, B. 2003. Risk of Malignancy Index in the preoperative evaluation of patients with adnexal masses. *Gynaecologic Oncology*, 90, 109-112.
- ANTON, C., CARVALHO, F. M., OLIVEIRA, E. I., MACIEL, G. A. R., BARACAT, E. C. & CARVALHO, J. P. 2012. A comparison of CA125, HE4, risk ovarian malignancy algorithm (ROMA), and risk malignancy index (RMI) for the classification of ovarian masses. *Clinics*, 67, 437-441.
- ARAB, M., YASERI, M., FARZANEH, M., MORIDI, A., TEHRANIAN, A. & SHEIBANI, K. 2010. The construction and validation of a new ovarian malignancy probability score (omps) for prediction of ovarian malignancy. *Iranian Journal of Cancer Prevention*, 3, 132-138.
- ARMITAGE, P., BERRY, G. & MATTHEWS, J. N. S. 2009. *Statistical methods in medical research*, Malden, Blackwell Science Ltd.
- ASLAM, N., TAILOR, A., LAWTON, F., CARR, J., SAVVAS, M. & JURKOVIC, D. 2000. Prospective evaluation of three different models for the pre-operative diagnosis of ovarian cancer. *British Journal of Obstetrics and Gynaecology*, 107, 1347-1353.
- BADER, W., BÖHMER, S., VAN LEEUWEN, P., HACKMANN, J., WESTHOF, G. & HATZMANN, W. 2000. Does texture analysis improve breast ultrasound precision? *Ultrasound in Obstetrics and Gynaecology*, 15, 311-316.
- BAILEY, E., FENNING, N., CHAMBERLAIN, S., DEVLIN, L., HOPKISSON, J. & TOMLINSON, M. 2007. Validation of sperm counting methods using limits of agreement. *Journal of Andrology*, 28, 364-373.
- BAILEY, J., TAILOR, A., NAIK, R., LOPES, A., GODFREY, K., HATEM, H. M. & MONAGHAN, J. 2006. Risk of malignancy index for referral of ovarian cancer cases to a tertiary centre: Does it identify the correct cases? *International Journal of Gynaecological Cancer*, 16, 30-34.
- BANKHEAD, C. R., COLLINS, C., STOKES-LAMPARD, H., ROSE, P., WILSON, S., CLEMENTS, A., MANT, D., KEHOE, S. T. & AUSTOKER, J. 2008. Identifying symptoms of ovarian cancer: A qualitative and quantitative study. *BJOG: An International Journal of Obstetrics and Gynaecology*, 115, 1008-1014.
- BARNETT, S. B. (ed.) 2000. *Ultrasound-induced heating and its biological consequences*, London: British Institute of Radiology.
- BARONI, M., FORTUNATO, P. & LA TORRE, A. 2007. Towards quantitative analysis of retinal features in optical coherence tomography. *Medical Engineering and Physics*, 29, 432-441.
- BARRETT, J., SHARP, D. J., STAPLEY, S., STABB, C. & HAMILTON, W. 2010. Pathways to the diagnosis of ovarian cancer in the UK: A cohort study in primary care. *BJOG: An International Journal of Obstetrics and Gynaecology*, 117, 610-614.

- BARUA, A., BITTERMAN, P., BAHR, J. M., BASU, S., SHEINER, E., BRADARIC, M. J., HALES, D. B., LUBORSKY, J. L. & ABRAMOWICZ, J. S. 2011. Contrast-enhanced sonography depicts spontaneous ovarian cancer at early stages in a preclinical animal model. *Journal of Ultrasound in Medicine*, 30, 333-345.
- BASSANI, D. G., MIRANDA, L. A. & GUSTAFSSON, A. 2007. Use of the limits of agreement approach in periodontology. *Oral health & preventive dentistry*, 5, 119-124.
- BASSET, O., SUN, Z., MESTAS, J. L. & GIMENEZ, G. 1993. Texture analysis of ultrasonic images of the prostate by means of co- occurrence matrices. *Ultrasonic Imaging*, 15, 218-237.
- BAST JR, R. C. 2003. Status of tumour markers in ovarian cancer screening. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 21, 200s-205s.
- BEEKMAN, R. & VISSER, L. H. 2004. High-resolution sonography of the peripheral nervous system - A review of the literature. *European Journal of Neurology*, 11, 305-314.
- BENACERRAF, B. R. 2008. Three-Dimensional Ultrasound in Gynaecology: Where is it Applicable? Available: http://www.gehealthcare.com/usen/ultrasound/products/msucme3d_gyn.html [Accessed 15 January 2013].
- BENEDET, J. L., BENDER, H., JONES 3RD, H., NGAN, H. Y. & PECORELLI, S. 2000. FIGO staging classifications and clinical practice guidelines in the management of gynaecologic cancers. FIGO Committee on Gynaecologic Oncology. *International Journal of Gynaecology and Obstetrics*, 70, 209-262.
- BERLANDA, N., FERRARI, M. M., MEZZOPANE, R., BOERO, V., GRIJUELA, B., FERRAZZI, E. & PARDI, G. 2002. Impact of a multiparameter, ultrasound-based triage on surgical management of adnexal masses. *Ultrasound in Obstetrics and Gynaecology*, 20, 181-185.
- BERRINO, F., DE ANGELIS, R., SANT, M., ROSSO, S., LASOTA, M. B., COEBERGH, J. W. & SANTAQUILANI, M. 2007. Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995-99: results of the EURO CARE-4 study. *Lancet Oncology*, 8, 773-783.
- BIAN, J., LI, B., KOU, X. J., LIU, T. Z. & MING, L. 2013. Clinical significance of combined detection of serum tumour markers in diagnosis of patients with ovarian cancer. *Asian Pacific Journal of Cancer Prevention*, 14, 6241-6243.
- BLACK, S. S., BUTLER, S. L., GOLDMAN, P. A. & SCROGGINS, M. J. 2007. Ovarian cancer symptom index: Possibilities for earlier detection. *Cancer*, 109, 167-169.

- BLAND, J. M. & ALTMAN, D. G. 2010. Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies*, 47, 931-936.
- BONILLA-MUSOLES, F., RAGA, F. & OSBORNE, N. G. 1995. Three-dimensional ultrasound evaluation of ovarian masses. *Gynaecologic Oncology*, 59, 129-135.
- BOOTE, E. J. 2003. Doppler US Techniques: Concepts of Blood Flow Detection and Flow Dynamics. *Radiographics*, 23, 1315-1327.
- BRAECKMAN, J., AUTIER, P., GARBAR, C., MARICHAL, M. P., SOVIANY, C., NIR, R., NIR, D., MICHELSEN, D., BLEIBERG, H., EGEVAD, L. & EMBERTON, M. 2008a. Computer-aided ultrasonography (HistoScanning): A novel technology for locating and characterizing prostate cancer. *BJU International*, 101, 293-298.
- BRAECKMAN, J., AUTIER, P., SOVIANY, C., NIR, R., NIR, D., MICHELSEN, D., TREURNICHT, K., JARMULOWICZ, M., BLEIBERG, H., GOVINDARAJU, S. & EMBERTON, M. 2008b. The accuracy of transrectal ultrasonography supplemented with computer-aided ultrasonography for detecting small prostate cancers. *BJU International*, 102, 1560-1565.
- BRIDAL, S. L., CORREAS, J. M., SAÏED, A. & LAUGIER, P. 2003. Milestones on the road to higher resolution, quantitative, and functional ultrasonic imaging. *Proceedings of the IEEE*, 91, 1543-1561.
- BROOKS, S. E. 1994. Preoperative Evaluation of Patients with Suspected Ovarian Cancer. *Gynaecologic Oncology*, 55, S80-S90.
- BROSKY, J. 2009. Toshiba ASQ delivers hard data on liver fibrosis. Available: www.european-hospital.com/en/article/6676-Toshiba_ASQ_delivers_hard_data_on_liver_fibrosis [Accessed 27-02-2014].
- BROWN, P. O. & PALMER, C. 2009. The preclinical natural history of serous ovarian cancer: Defining the target for early detection. *PLoS Medicine*, 6.
- BROWNE, J. E., WATSON, A. J., GIBSON, N. M., DUDLEY, N. J. & ELLIOTT, A. T. 2004. Objective measurements of image quality. *Ultrasound in Medicine & Biology*, 30, 229-237.
- BRUTON, A., CONWAY, J. H. & HOLGATE, S. T. 2000. Reliability: What is it, and how is it measured? *Physiotherapy*, 86, 94-99.
- BUCKSHEE, K., TEMSU, I., BHATLA, N. & DEKA, D. 1998. Pelvic examination, transvaginal ultrasound and transvaginal colour Doppler sonography as predictors of ovarian cancer. *International Journal of Gynaecology and Obstetrics*, 61, 51-57.
- BUYS, S. S., PARTRIDGE, E., BLACK, A., JOHNSON, C. C., LAMERATO, L., ISAACS, C., REDING, D. J., GREENLEE, R. T., YOKOCHI, L. A., KESSEL, B., CRAWFORD, E. D., CHURCH, T. R., ANDRIOLE, G. L., WEISSFELD, J. L., FOUAD, M. N., CHIA, D., O'BRIEN, B., RAGARD, L. R., CLAPP, J. D.,

- RATHMELL, J. M., RILEY, T. L., HARTGE, P., PINSKY, P. F., ZHU, C. S., IZMIRLIAN, G., KRAMER, B. S., MILLER, A. B., XU, J. L., PROROK, P. C., GOHAGAN, J. K. & BERG, C. D. 2011. Effect of screening on ovarian cancer mortality: The Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening randomized controlled trial. *JAMA - Journal of the American Medical Association*, 305, 2295-2302.
- CALONGE, N., ALLAN, J. D., BERG, A. O., FRAME, P. S., GARCIA, J., GORDIS, L., GREGORY, K. D., HARRIS, R., JOHNSON, M. S., KLEIN, J. D., LOVELAND-CHERRY, C., MOYER, V. A., OCKENE, J. K., PETITTI, D. B., SIU, A. L., TEUTSCH, S. M. & YAWN, B. P. 2004. Screening for ovarian cancer: Recommendation statement. *Annals of Family Medicine*, 2, 260-262.
- CAMPBELL, S., BHAN, V., ROYSTON, P., WHITEHEAD, M. I. & COLLINS, W. P. 1989. Transabdominal ultrasound screening for early ovarian cancer. *British Medical Journal*, 299, 1363-1367.
- CANCER RESEARCH, U. 2011. Ovarian cancer - UK incidence statistics. Available: <http://info.cancerresearchuk.org/cancerstats/types/ovary/incidence> /. [Accessed 31 October 2011].
- CANCER RESEARCH, U. 2012. Statistics and outlook for ovarian cancer [Online]. Available: <http://www.cancerresearchuk.org/cancer-help/type/ovarian-cancer/treatment/statistics-and-outlook-for-ovarian-cancer#overall> [Accessed 17 January 2013].
- CANCER RESEARCH, U. 2015. Ovarian Cancer Incidence [Online]. Available: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer#heading-Zero> [Accessed 20-10-2015].
- CASS, I. 2009. The search for meaning - Symptoms and transvaginal sonography screening for ovarian cancer: Silent no more. *Cancer*, 115, 3606-3608.
- CASTELLANO, G., BONILHA, L., LI, L. M. & CENDES, F. 2004. Texture analysis of medical images. *Clinical Radiology*, 59, 1061-1069.
- CHAN, H. P., DOI, K., GALHOTRA, S., VYBORNY, C. J., MACMAHON, H., JOKICH, P. M. & GIGER, M. L. 1987. Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography. *Medical Physics*, 14, 538-548.
- CHAN, K. K. & SELMAN, T. J. 2006. Testing for ovarian cancer. *Best Practice and Research: Clinical Obstetrics and Gynaecology*, 20, 977-983.
- CHAN, K. L. & MCCARTY, K. 1990. Aspects of the statistical texture analysis of the medical ultrasound images. *IEE Colloquium on Ultrasound Instrumentation*, 1/3-3/3.
- CHEN, D.-R., CHANG, R.-F., KUO, W.-J., CHEN, M.-C. & HUANG, Y. U.-L. 2002. Diagnosis of breast tumours with sonographic texture analysis using wavelet

transform and neural networks. *Ultrasound in Medicine & Biology*, 28, 1301-1310.

CHEN, S.-J., LIN, C.-H., CHANG, C.-Y., CHANG, K.-Y., HO, H.-C., HSIAO, S.-H., LIN, C.-W., TZENG, J.-E., CHEN, Y.-T. & TSAI, H.-M. 2012. Characterizing the major sonographic textural difference between metastatic and common benign lymph nodes using support vector machine with histopathologic correlation. *Clinical Imaging*, 36, 353-359.e2.

CHIA, Y. N., MARSDEN, D. E., ROBERTSON, G. & HACKER, N. F. 2008. Triage of ovarian masses. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 48, 322-328.

CHU, C. S. & RUBIN, S. C. 2006. Screening for ovarian cancer in the general population. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 20, 307-320.

CLARK, T. G., STEWART, M., RYE, T., SMYTH, J. F. & GOURLEY, C. 2007. Validation of a new prognostic index for advanced epithelial ovarian cancer: Results from its application to a UK-based cohort [8]. *Journal of Clinical Oncology*, 25, 5669-5670.

CLARKE-PEARSON, D. L. 2009. Screening for ovarian cancer. *New England Journal of Medicine*, 361, 170-177.

CLARKE, S. E., GRIMSHAW, R., RITTENBERG, P., KIESER, K. & BENTLEY, J. 2009. Risk of malignancy index in the evaluation of patients with adnexal masses. *Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstétrique et gynécologie du Canada : JOGC*, 31, 440-445.

COHEN, L. S., ESCOBAR, P. F., SCHARM, C., GLIMCO, B. & FISHMAN, D. A. 2001. Three-Dimensional Power Doppler Ultrasound Improves the Diagnostic Accuracy for Ovarian Cancer Prediction. *Gynaecologic Oncology*, 82, 40-48.

COLEMAN, D. P., RAKEBRANDT, F., PUGH, N. D., CRAWFORD, D. C. & WOODCOCK, J. P. 2005. Development and validation of an in vivo analysis tool to identify changes in carotid plaque tissue types in serial 3-D ultrasound scans. *Ultrasound in Medicine and Biology*, 31, 329-335.

COLLEWET, G., STRZELECKI, M. & MARIETTE, F. 2004. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magnetic Resonance Imaging*, 22, 81-91.

CRAGUN, J. M. 2011. Screening for ovarian cancer. *Cancer Control*, 18, 16-21.

CRAWFORD, I. M. 1990. Marketing research and information systems. *Agriculture and Consumer Protection*. [http://www.fao.org/docrep/w3241E/w3241e05](http://www.fao.org/docrep/w3241E/w3241e05.htm) .htm: FAO Corporate Document Repository.

CRUM, C. P. 2004. *The Female Genital Tract*, Philadelphia, Pennsylvania, Elsevier Inc.

- DAI, S. Y., HATA, K., INUBASHIRI, E., KANENISHI, K., SHIOTA, A., OHNO, M., YAMAMOTO, Y., NISHIYAMA, Y., OHKAWA, M. & HATA, T. 2008. Does three-dimensional power Doppler ultrasound improve the diagnostic accuracy for the prediction of adnexal malignancy? *Journal of Obstetrics and Gynaecology Research*, 34, 364-370.
- DAS, A., NGUYEN, C. C., LI, F. & LI, B. 2008. Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue. *Gastrointestinal Endoscopy*, 67, 861-867.
- DE KANT, J. V. 2011. A guide to Acoustic Structure Quantification ASQ Apolio XG V5 Rev 2. In: TOSHIBA (ed.). Toshiba leading innovation.
- DE VET, H. C. W., TERWEE, C. B., KNOL, D. L. & BOUTER, L. M. 2006. When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, 59, 1033-1039.
- DECRUZE, S. B. & KIRWAN, J. M. 2006. Ovarian cancer. *Current Obstetrics & Gynaecology*, 16, 161-167.
- DEPRIEST, P. D. & DESIMONE, C. P. 2003. Ultrasound screening for the early detection of ovarian cancer. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 21, 194s-199s.
- DEPRIEST, P. D., GALLION, H. H., PAVLIK, E. J., KRYSCIO, R. J. & VAN NAGELL, J. R. 1997. Transvaginal Sonography as a Screening Method for the Detection of Early Ovarian Cancer. *Gynaecologic Oncology*, 65, 408-414.
- DEPRIEST, P. D., SHENSON, D., FRIED, A., HUNTER, J. E., ANDREWS, S. J., GALLION, H. H., PAVLIK, E. J., KRYSCIO, R. J. & VAN NAGELL JR, J. R. 1993. A morphology index based on sonographic findings in ovarian cancer. *Gynaecologic Oncology*, 51, 7-11.
- DEPRIEST, P. D. & VAN NAGELL JR, J. R. 1996. Ovarian cancer screening. *Current Problems in Obstetrics, Gynaecology and Fertility*, 19, 99-109.
- DEPRIEST, P. D., VARNER, E., POWELL, J., FRIED, A., PULS, L., HIGGINS, R., SHENSON, D., KRYSCIO, R., HUNTER, J. E., ANDREWS, S. J. & VAN NAGELL JR, J. R. 1994. The efficacy of a sonographic morphology index in identifying ovarian cancer: A multi-institutional investigation. *Gynaecologic Oncology*, 55, 174-178.
- DERCHI, L. E., SERAFINI, G., GANDOLFO, N., GANDOLFO, N. G. & MARTINOLI, C. 2001. Ultrasound in Gynaecology. *European Radiology*, 11, 2137-2155.
- DIAMOND, J., ANDERSON, N. H., BARTELS, P. H., MONTIRONI, R. & HAMILTON, P. W. 2004. The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia. *Human Pathology*, 35, 1121-1131.

- DODGE, J. E., COVENS, A. L., LACCHETTI, C., ELIT, L. M., LE, T., DEVRIES-ABOUD, M. & FUNG-KEE-FUNG, M. 2012. Preoperative identification of a suspicious adnexal mass: A systematic review and meta-analysis. *Gynaecologic Oncology*, 126, 157-166.
- DOI, K. 2007. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31, 198-211.
- DONG, L., CUI, H., LI, X. P., SUN, L. F., CHANG, X. H., LIANG, X. D. & ZHU, H. L. 2008. Clinical value of serum CA19-9, CA125 and CP2 in mucinous ovarian tumour: a retrospective study of 273 patients. *Zhonghua fu chan ke za zhi*, 43, 5-8.
- DUDELY, N. 2010. *B-mode measurement*, Cambridge, Cambridge University Press.
- DUFFY, M. J., BONFRER, J. M., KULPA, J., RUSTIN, G. J. S., SOLETORMOS, G., TORRE, G. C., TUXEN, M. K. & ZWIRNER, M. 2005. CA125 in ovarian cancer: European Group on Tumour Markers guidelines for clinical use. *International Journal of Gynaecological Cancer*, 15, 679-691.
- DUTTA, S., WANG, F. Q., FLEISCHER, A. C. & FISHMAN, D. A. 2010a. New frontiers for ovarian cancer risk evaluation: Proteomics and contrast-enhanced ultrasound. *American Journal of Roentgenology*, 194, 349-354.
- DUTTA, S., WANG, F. Q., PHALEN, A. & FISHMAN, D. A. 2010b. Biomarkers for ovarian cancer detection and therapy. *Cancer Biology and Therapy*, 9, 666-675.
- EBELL, M. H. online. Chapter 4, diagnostic tests. Available: <http://ebp.uga.edu/courses/Chapter%204%20-%20Diagnosis%20I%20-%20ROC%20curves.html> [Accessed 29-05-2015].
- EDWARDS, B. K., BROWN, M. L., WINGO, P. A., HOWE, H. L., WARD, E., RIES, L. A. G., SCHRAG, D., JAMISON, P. M., JEMAL, A., WU, X. C., FRIEDMAN, C., HARLAN, L., WARREN, J., ANDERSON, R. N. & PICKLE, L. W. 2005. Annual report to the nation on the status of cancer, 1975-2002, featuring population-based trends in cancer treatment. *Journal of the National Cancer Institute*, 97, 1407-1427.
- EKERHOVD, E., WIENERROITH, H., STAUDACH, A. & GRANBERG, S. 2001. Preoperative assessment of unilocular adnexal cysts by transvaginal ultrasonography: A comparison between ultrasonographic morphologic imaging and histopathologic diagnosis. *American Journal of Obstetrics and Gynaecology*, 184, 48-54.
- ENAKPENE, C. A., OMIGBODUN, A. O., GOECKE, T. W., ODUKOGBE, A. T. & BECKMANN, M. W. 2009. Preoperative evaluation and triage of women with suspicious adnexal masses using risk of malignancy index. *Journal of Obstetrics and Gynaecology Research*, 35, 131-138.
- ENGELAND, A., TRETTLI, S. & BJÄRGE, T. 2003. Height, body mass index, and ovarian cancer: A follow-up of 1.1 million Norwegian women. *Journal of the National Cancer Institute*, 95, 1244-1248.

- ERDOGAN, N., Ã–ZÄŞELIK, B., SERIN, I. S., AKGÄ¼N, M. & Ã–ZTÄ¼RK, F. 2005. Doppler ultrasound assessment and serum cancer antigen 125 in the diagnosis of ovarian tumours. *International Journal of Gynaecology and Obstetrics*, 91, 146-150.
- EXACOUSTOS, C., ROMANINI, M. E., RINALDO, D., AMOROSO, C., SZABOLCS, B., ZUPI, E. & ARDUINI, D. 2005. Preoperative sonographic features of borderline ovarian tumours. *Ultrasound in Obstetrics and Gynaecology*, 25, 50-59.
- FASCHINGBAUER, F., BENZ, M., HÄBERLE, L., GOECKE, T. W., BECKMANN, M. W., RENNER, S., MÜLLER, A., WITTENBERG, T. & MÜNZENMAYER, C. 2012. Subjective assessment of ovarian masses using pattern recognition: the impact of experience on diagnostic performance and interobserver variability. *Archives of Gynaecology and Obstetrics*, 1-7.
- FERLAY, J., BRAY, F., PISANI, P. & PARKIN, D. 2004. GLOBOCAN 2002: cancer incidence, mortality and prevalence worldwide. IARC Cancer Base No. 5, version 2.0.
- FERLAY, J., PARKIN, D. M. & STELIAROVA-FOUCHER, E. 2010. Estimates of cancer incidence and mortality in Europe in 2008. *European Journal of Cancer*, 46, 765-781.
- FERRAZZI, E., ZANETTA, G., DORDONI, D., BERLANDA, N., MEZZOPANE, R. & LISSONI, G. 1997. Transvaginal ultrasonographic characterization of ovarian masses: Comparison of five scoring systems in a multicentre study. *Ultrasound in Obstetrics and Gynaecology*, 10, 192-197.
- FIROOZABADI, R. D., ZARCHI, M. K., MANSURIAN, H. R., MOGHADAM, B. R., TEIMOORI, S. & NASERI, A. 2011. Evaluation of diagnostic value of CT scan, physical examination and ultrasound based on pathological findings in patients with pelvic masses. *Asian Pacific Journal of Cancer Prevention*, 12, 1745-1747.
- FISCHEROVA, D. 2011. Ultrasound scanning of the pelvis and abdomen for staging of gynaecological tumours: A review. *Ultrasound in Obstetrics and Gynaecology*, 38, 246-266.
- FISHMAN, D. A. & AL., E. 2001. The role of ultrasound in detecting early ovarian carcinoma: The National Ovarian Cancer Early Detection Program. *Medica Mundi*, 45, 42- 47.
- FITCH, M. I., GRAY, R. E., COVENS, A., THOMAS, G., FRANSSEN, E., DEPETRILLO, D. & ROSEN, B. 1999. Gynecologists' perspectives regarding ovarian cancer. *Cancer prevention & control : CPC = Prévention & contrôle en cancérologie : PCC*, 3, 68-76.
- FLEISCHER, A. C. 2005. Recent advances in the sonographic assessment of vascularity and blood flow in gynaecologic conditions. *American Journal of Obstetrics and Gynaecology*, 193, 294-301.

- FLEISCHER, A. C., LYSHCHIK, A., HIRARI, M., MOORE, R. D., ABRAMSON, R. G. & FISHMAN, D. A. 2012. Early detection of ovarian cancer with conventional and contrast-enhanced transvaginal sonography: Recent advances and potential improvements. *Journal of Oncology*.
- FRY, T. 2002. Evaluation of the Repeatability of Residual Stress Measurements Using X-Ray Diffraction, Middlesex, National Physical Laboratory.
- FUNG, E. T., HOGDALL, C., EENGELHOLM, S. A., PETRI, A. L. & NEDERGAARD, L. 2006. Novel biomarkers to aid in the differential diagnosis of a pelvic mass, Santa, Monica, CA.
- GADDUCCI, A., COSIO, S., CARPI, A., NICOLINI, A. & GENAZZANI, A. R. 2004. Serum tumour markers in the management of ovarian, endometrial and cervical cancer. *Biomedicine and Pharmacotherapy*, 58, 24-38.
- GAGNON, A. & YE, B. 2008. Discovery and application of protein biomarkers for ovarian cancer. *Current Opinion in Obstetrics and Gynaecology*, 20, 9-13.
- GAO, S., PENG, Y., GUO, H., LIU, W., GAO, T., XU, Y. & TANG, X. 2014. Texture analysis and classification of ultrasound liver images. *Bio-Medical Materials and Engineering*, 24, 1209-1216.
- GARRA, B. S., KRASNER, B. H., HORII, S. C., ASCHER, S., MUN, S. K. & ZEMAN, R. K. 1993. Improving the Distinction Between Benign and Malignant Breast Lesions: The Value of Sonographic Texture Analysis. *Ultrasonic Imaging*, 15, 267-285.
- GENTRY-MAHARAJ, A. & MENON, U. 2012. Screening for ovarian cancer in the general population. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 26, 243-256.
- GEOMINI, P., KRUITWAGEN, R., BREMER, G. L., CNOSSEN, J. & MOL, B. W. J. 2009. The accuracy of risk scores in predicting ovarian malignancy: A systematic review. *Obstetrics and Gynaecology*, 113, 384-394.
- GEORGE, D., GEORGE, P. & VASILLIS, M. 2013. The effect of four-week interrupted intervention Whole-Body Vibration program on hamstring's flexibility. *Journal of Physical Education and Sport*, 13, 517-521.
- GEORGE, L. 2006. *Ultrasound Physics. the Westmead TOE Manual* Sydney: Westmead Hospital.
- GIGER, M. L., CHAN, H. P. & BOONE, J. 2008. Anniversary paper: History and status of CAD and quantitative image analysis: The role of Medical Physics and AAPM. *Medical Physics*, 35, 5799-5820.
- GIGER, M. L., DOI, K. & MACMAHON, H. 1987. Computerized detection of lung nodules in digital chest radiographs *Proc. SPIE*, 767, 384-386.

- GIVENS, V., MITCHELL, G., HARRAWAY-SMITH, C., REDDY, A. & MANESS, D. L. 2009. Diagnosis and management of adnexal masses. *American Family Physician*, 80.
- GLETSOS, M., MOUGIAKAKOU, S. G., MATSOPOULOS, G. K., NIKITA, K. S., NIKITA, A. S. & KELEKIS, D. 2003. A Computer-Aided Diagnostic System to Characterize CT Focal Liver Lesions: Design and Optimization of a Neural Network Classifier. *IEEE Transactions on Information Technology in Biomedicine*, 7, 153-162.
- GOFF, B. A., LOWE, K. A., KANE, J. C., ROBERTSON, M. D., GAUL, M. A. & ANDERSEN, M. R. 2012. Symptom triggered screening for ovarian cancer: A pilot study of feasibility and acceptability. *Gynaecologic Oncology*, 124, 230-235.
- GOFF, B. A., MANDEL, L., MUNTZ, H. G. & MELANCON, C. H. 2000. Ovarian carcinoma diagnosis: Results of a national ovarian cancer survey. *Cancer*, 89, 2068-2075.
- GOFF, B. A., MANDEL, L. S., DRESCHER, C. W., URBAN, N., GOUGH, S., SCHURMAN, K. M., PATRAS, J., MAHONY, B. S. & ROBYN ANDERSEN, M. 2007. Development of an ovarian cancer symptom index: Possibilities for earlier detection. *Cancer*, 109, 221-227.
- GORELIK, E., LANDSITTEL, D. P., MARRANGONI, A. M., MODUGNO, F., VELIKOKHATNAYA, L., WINANS, M. T., BIGBEE, W. L., HERBERMAN, R. B. & LOKSHIN, A. E. 2005. Multiplexed immunobead-based cytokine profiling for early detection of ovarian cancer. *Cancer Epidemiology Biomarkers and Prevention*, 14, 981-987.
- GOSSAGE, K. W., TKACZYK, T. S., RODRIGUEZ, J. J. & BARTON, J. K. Year. Texture analysis for tissue classification of optical coherence tomography images. In, 2003. 109-117.
- GRAB, D., FLOCK, F., STOHR, I., RIEBER, A., NUSSLE, S., BRAMBS, H. J., FENCHEL, S., RESKE, S. N. & KREIENBERG, R. 2000. Premenopausal adnexal masses: Diagnostic accuracy of sonography, magnetic resonance imaging, and positron emission tomography. *Diagnostische wertigkeit von sonographie, kernspintomographie und positronenemissions-tomographie zur beurteilung tier dignitat von adnextumouren bei pramenopausalen patientinnen*, 60, 544-552.
- GRAMELLINI, D., FIENI, S., SANAPO, L., CASILLA, G., VERROTTI, C. & NARDELLI, G. B. 2008. Diagnostic accuracy of IOTA ultrasound morphology in the hands of less experienced sonographers. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 48, 195-201.
- GRANBERG, S., WIKLAND, M. & JANSSON, I. 1989. Macroscopic characterization of ovarian tumours and the relation to the histological diagnosis: Criteria to be used for ultrasound evaluation. *Gynaecologic Oncology*, 35, 139-144.

GREEN, A., PURDIE, D., BAIN, C., SISKIND, V., RUSSELL, P., QUINN, M. & WARD, B. 1997. Tubal sterilisation, hysterectomy and decreased risk of ovarian cancer. *International Journal of Cancer*, 71, 948-951.

GREENLEE, R. T., KESSEL, B., WILLIAMS, C. R., RILEY, T. L., RAGARD, L. R., HARTGE, P., BUYS, S. S., PARTRIDGE, E. E. & REDING, D. J. 2010. Prevalence, incidence, and natural history of simple ovarian cysts among women >55 years old in a large cancer screening trial. *American Journal of Obstetrics and Gynaecology*, 202, 373.e1-373.e9.

GREISER, C. M., GREISER, E. M. & DÅRREN, M. 2007. Menopausal hormone therapy and risk of ovarian cancer: Systematic review and meta-analysis. *Human Reproduction Update*, 13, 453-463.

GUERRIERO, S., AJOSSA, S., GARAU, N., PIRAS, B., PAOLETTI, A. M. & MELIS, G. B. 2005. Ultrasonography and colour Doppler-based triage for adnexal masses to provide the most appropriate surgical approach. *American Journal of Obstetrics and Gynaecology*, 192, 401-406.

GUERRIERO, S., AJOSSA, S., GERADA, M., MELIS, G. B. & ALCÁZAR, J. L. 2006. Pattern recognition and descriptive sonographic scoring in the diagnosis of ovarian cancer [1]. *Journal of Ultrasound in Medicine*, 25, 558-559.

GUERRIERO, S., AJOSSA, S., RISALVATO, A., LAI, M. P., MAIS, V., ANGIOLUCCI, M. & MELIS, G. B. 1998. Diagnosis of adnexal malignancies by using colour Doppler energy imaging as a secondary test in persistent masses. *Ultrasound in Obstetrics and Gynaecology*, 11, 277-282.

GUERRIERO, S., ALCAZAR, J. L., AJOSSA, S., LAI, M. P., ERRASTI, T., MALLARINI, G. & MELIS, G. B. 2001. Comparison of conventional colour Doppler imaging and power Doppler imaging for the diagnosis of ovarian cancer: Results of a European study. *Gynaecologic Oncology*, 83, 299-304.

GUERRIERO, S., ALCAZAR, J. L., PASCUAL, M. A., AJOSSA, S., GRAUPERA, B., HERETER, L. & MELIS, G. B. 2011. The diagnosis of ovarian cancer: Is colour Doppler imaging reproducible and accurate in examiners with different degrees of experience? *Journal of Women's Health*, 20, 273-277.

GUNNELL, D. 2001. Commentary: Early insights into height, leg length, proportionate growth and health. *International Journal of Epidemiology*, 30, 221-222.

GUY, C. (ed.) 2008. *An Introduction to the principles of medical imaging*, London: Imperial College.

HAIYAN, H. & MIN, D. 2011. Ultrasound operators' confidence influences diagnosis of ovarian tumours - A study in China. *Asian Pacific Journal of Cancer Prevention*, 12, 1275-1277.

HÅKANSSON, F., HØGDALL, E. V. S., NEDERGAARD, L., LUNDEVALL, L., ENGELHOLM, S. A., PEDERSEN, A. T., HARTWELL, D. & HØGDALL, C. 2012.

Risk of malignancy index used as a diagnostic tool in a tertiary centre for patients with a pelvic mass. *Acta Obstetrica et Gynaecologica Scandinavica*, 91, 496-502.

HAMID, B., 2011. The Reliability of B-mode transvaginal probe image for the quantitative texture analysis and the dependence of extracted features on region of interest size for ovarian cancer detection. PhD, Cardiff University.

HAMILTON, W. 2012. Computer assisted diagnosis of ovarian cancer in primary care. *BMJ (Online)*, 344.

HAMILTON, W., ROUND, A. & SHARP, D. 2009. Ovarian cancer: Not a silent killer. *BMJ*, 339, 123.

HANGIANDREOU, N. J. 2003. B-mode US: Basic concepts and new technology. *Radiographics*, 23, 1019- 1033.

HANKINSON, S. E., HUNTER, D. J., COLDITZ, G. A., WILLETT, W. C., STAMPFER, M. J., ROSNER, B., HENNEKENS, C. H. & SPEIZER, F. E. 1993. Tubal ligation, hysterectomy, and risk of ovarian cancer: A prospective study. *Journal of the American Medical Association*, 270, 2813-2818.

HARALICK, R. M., SHANMUGAM, K. & DINSTEN, I. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, smc 3, 610-621.

HARRIS, L. L. 2002. Ovarian cancer: Screening for early detection. *American Journal of Nursing*, 102, 46-52.

HARRISON, L. C. V., RAUNIO, M., HOLLI, K. K., LUUKKAALA, T., SAVIO, S., ELOVAARA, I., SOIMAKALLIO, S., ESKOLA, H. J. & DASTIDAR, P. 2010. MRI Texture Analysis in Multiple Sclerosis: Toward a Clinical Analysis Protocol. *Academic Radiology*, 17, 696-707.

HARRY, V. N., NARAYANSINGH, G. V. & PARKIN, D. E. 2009. The risk of malignancy index for ovarian tumours in northeast Scotland - A population based study. *Scottish Medical Journal*, 54, 21-23.

HARTMAN, C. A., JULIATO, C. R. T., SARIAN, L. O., TOLEDO, M. C., JALES, R. M., MORAIS, S. S., PITTA, D. D., MARUSSI, E. F. & DERCHAIN, S. 2012. Ultrasound criteria and CA 125 as predictive variables of ovarian cancer in women with adnexal tumours. *Ultrasound in Obstetrics and Gynaecology*, 40, 360-366.

HASSEN, K., GHOSAIN, M. A., ROUSSET, P., SCIOT, C., HUGOL, D., BADDOURA, R., VADROT, D. & BUY, J. N. 2011. Characterization of papillary projections in benign versus borderline and malignant ovarian masses on conventional and colour doppler ultrasound. *American Journal of Roentgenology*, 196, 1444-1449.

HAVRILESKY, L. J., SANDERS, G. D., KULASINGAM, S. & MYERS, E. R. 2008. Reducing ovarian cancer mortality through screening: Is it possible, and can we afford it? *Gynaecologic Oncology*, 111, 179-187.

- HE, Z., HUSTON, D. R., GRIMM, S., JANNICKY, E., GARRA, B. S., WAGNER, R. F. & WEAR, K. A. Year. Dependence of tissue characterization features on Region of Interest (ROI) size: Studies on phantoms and simulations. In, 2004. 2082-2085.
- HECHT, J. L., KOTSOPOULOS, J., HANKINSON, S. E. & TWOROGGER, S. S. 2009. Relationship between epidemiologic risk factors and hormone receptor expression in ovarian cancer: Results from the nurses' health study. *Cancer Epidemiology Biomarkers and Prevention*, 18, 1624-1630.
- HELLSTROM, I. & HELLSTROM, K. E. 2008. SMRP and HE4 as biomarkers for ovarian carcinoma when used alone and in combination with CA125 and/or each other.
- HENSLEY, M. L. 2010. A step forward for two-step screening for ovarian cancer. *Journal of Clinical Oncology*, 28, 2128-2130.
- HENSLEY, M. L., CASTIEL, M. & ROBSON, M. E. 2000. Screening for ovarian cancer: What we know, what we need to know. *ONCOLOGY*, 14, 1601-1607.
- HERLIDOU-MÊME, S., CONSTANS, J. M., CARSIN, B., OLIVIE, D., ELIAT, P. A., NADAL-DESBARATS, L., GONDRY, C., LE RUMEUR, E., IDY-PERETTI, I. & DE CERTAINES, J. D. 2003. MRI texture analysis on texture test objects, normal brain and intracranial tumours. *Magnetic Resonance Imaging*, 21, 989-993.
- HOLLI, K. K., HARRISON, L., DASTIDAR, P., WÄLJAS, M., LIIMATAINEN, S., LUUKKAALA, T., ÖHMAN, J., SOIMAKALLIO, S. & ESKOLA, H. 2010. Texture analysis of MR images of patients with Mild Traumatic Brain Injury. *BMC Medical Imaging*, 10.
- HORSKIN, P. 2010. *Principle of Doppler Ultrasound*, Cambridge: Cambridge University Press.
- HOSSAIN, F., KARIM, M. N., RAHMAN, S. M. M., KHAN, N., SIDDIQUI, M. & HUSSAIN, R. 2010. Preoperative detection of ovarian cancer by colour Doppler ultrasonography and CA 125. *Bangladesh Medical Research Council Bulletin*, 36, 68-73.
- HUANG, Y. L., CHEN, D. R., JIANG, Y. R., KUO, S. J., WU, H. K. & MOON, W. K. 2008. Computer-aided diagnosis using morphological features for classifying breast lesions on ultrasound. *Ultrasound in Obstetrics and Gynaecology*, 32, 565-572.
- HUBER, S., DANES, J., ZUNA, I., TEUBNER, J., MEDL, M. & DELORME, S. 2000. Relevance of sonographic B-mode criteria and computer-aided ultrasonic tissue characterization in differential/diagnosis of solid breast masses. *Ultrasound in Medicine & Biology*, 26, 1243-1252.
- HUCHON, C., METZGER, U., BATS, A. S., BENSALD, C., CHATELLIER, G., AZIZI, M., LEFRÈRE-BELDA, M. A., DUJARDIN, A., BERNARD, J. P. & LÉCURU, F. 2012. Value of three-dimensional contrast-enhanced power Doppler ultrasound for characterizing adnexal masses. *Journal of Obstetrics and Gynaecology Research*, 38, 832-840.

- HUNG, N. T. 2010. Lesson preparation of Ultrasonic diagnosis. nguyenthienhung.com
- HUYNEN, A. L., GIESEN, R. J. B., DE LA ROSETTE, J. J. M. C. H., AARNINK, R. G., DEBRUYNE, F. M. J. & WIJKSTRA, H. 1994. Analysis of ultrasonographic prostate images for the detection of prostatic carcinoma: The Automated Urologic Diagnostic Expert system. *Ultrasound in Medicine and Biology*, 20, 1-10.
- INTERNATIONAL AGENCY FOR RESEARCH ON CANCER, E. 2012. Ovarian cancer: Estimated incidence, mortality & prevalence, 2012 [Online]. World Health Organization Available: <http://eucancer.iarc.fr/EUCAN/CancerOne.aspx?Cancer=27&Gender=2> [Accessed 10-02-2016].
- ITAKURA, T., KIKKAWA, F., KAJIYAMA, H., MITSUI, T., KAWAI, M. & MIZUTANI, S. 2003. Doppler flow and arterial location in ovarian tumours. *International Journal of Gynaecology and Obstetrics*, 83, 277-283.
- IYER, V. R. & LEE, S. I. 2010. MRI, CT, and PET/CT for ovarian cancer detection and adnexal lesion characterization. *American Journal of Roentgenology*, 194, 311-321.
- JACOBS, I., ORAM, D., FAIRBANKS, J., TURNER, J., FROST, C. & GRUDZINSKAS, J. G. 1990. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *British Journal of Obstetrics and Gynaecology*, 97, 922-929.
- JACOBS, I., STABILE, I., BRIDGES, J., KEMSLEY, P., REYNOLDS, C., GRUDZINSKAS, J. & ORAM, D. 1988. Multimodal approach to screening for ovarian cancer. *Lancet*, 1, 268-271.
- JACOBS, I. J. & MENON, U. 2004. Progress and challenges in screening for early detection of ovarian cancer. *Molecular and Cellular Proteomics*, 3, 355-366.
- JELOVAC, D. & ARMSTRONG, D. K. 2011. Recent progress in the diagnosis and treatment of ovarian cancer. *CA Cancer Journal for Clinicians*, 61, 183-203.
- JEMAL, A., SIEGEL, R., WARD, E., HAO, Y., XU, J. & THUN, M. J. 2009. Cancer statistics, 2009. *CA Cancer Journal for Clinicians*, 59, 225-249.
- JEONG, Y. Y., OUTWATER, E. K. & KANG, H. K. 2000. From the RSNA refresher courses: Imaging evaluation of ovarian masses. *Radiographics*, 20, 1445-1470.
- JI, Q., ENGEL, J. & CRAINE, E. 2000. Texture analysis for classification of cervix lesions. *IEEE Transactions on Medical Imaging*, 19, 1144-1149.
- JOKUBKIENE, L., SLADKEVICIUS, P. & VALENTIN, L. 2007. Does three-dimensional power Doppler ultrasound help in discrimination between benign and malignant ovarian masses? *Ultrasound in Obstetrics and Gynaecology*, 29, 215-225.

- KADER ALI MOHAN, G. R., JAABACK, K., PROIETTO, A., ROBERTSON, R. & ANGSTETRA, D. 2010. Risk Malignancy Index (RMI) in patients with abnormal pelvic mass: Comparing RMI 1, 2 and 3 in an Australian population. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 50, 77-80.
- KAIJSER, J., BOURNE, T., VALENTIN, L., SAYASNEH, A., VAN HOLSBEKE, C., VERGOTE, I., TESTA, A. C., FRANCHI, D., VAN CALSTER, B. & TIMMERMAN, D. 2013. Improving strategies for diagnosing ovarian cancer: A summary of the International Ovarian Tumour Analysis (IOTA) studies. *Ultrasound in Obstetrics and Gynaecology*, 41, 9-20.
- KAIJSER, J., SAYASNEH, A., VAN HOORDE, K., GHAEM-MAGHAMI, S., BOURNE, T., TIMMERMAN, D. & VAN CALSTER, B. 2014a. Presurgical diagnosis of adnexal tumours using mathematical models and scoring systems: A systematic review and meta-analysis. *Human Reproduction Update*, 20, 449-462.
- KAIJSER, J., VANDECAVEYE, V., DEROOSE, C. M., ROCKALL, A., THOMASSIN-NAGGARA, I., BOURNE, T. & TIMMERMAN, D. 2014b. Imaging techniques for the pre-surgical diagnosis of adnexal tumours. *Best Practice and Research: Clinical Obstetrics and Gynaecology*, 28, 683-695.
- KALGHATGI-KULKARNI, K. & KUSHTAGI, P. 2008. Ovarian crescent sign and sonomorphological indices in preoperative determination of malignancy in adnexal masses. *Indian Journal of Medical Sciences*, 62, 477-483.
- KARAHALIOU, A., SKIADOPOULOS, S., BONIATIS, I., SAKELLAROPOULOS, P., LIKAKI, E., PANAYIOTAKIS, G. & COSTARIDOU, L. 2007. Texture analysis of tissue surrounding microcalcifications on mammograms for breast cancer diagnosis. *British Journal of Radiology*, 80, 648-656.
- KARANICOLAS, P. J., BHANDARI, M., KREDER, H., MORONI, A., RICHARDSON, M., WALTER, S. D., NORMAN, G. R. & GUYATT, G. H. 2009. Evaluating agreement: Conducting a reliability study. *Journal of Bone and Joint Surgery - Series A*, 91, 99-106.
- KARLAN, B. Y. & PLATT, L. D. 1995. Ovarian cancer screening: The role of ultrasound in early detection. *Cancer*, 76, 2011-2015.
- KEYS, A., FIDANZA, F., KARVONEN, M. J., KIMURA, N. & TAYLOR, H. L. 1972. Indices of relative weight and obesity. *Journal of Chronic Diseases*, 25, 329-343.
- KINKEL, K., HRICAK, H., LU, Y., TSUDA, K. & FILLY, R. A. 2000. US characterization of ovarian masses: A meta-analysis. *Radiology*, 217, 803-811.
- KLEINBAUM, D. 2010. *Logistic Regression: A self-Learning text*, Springer.
- KOBAYASHI, H., YAMADA, Y., SADO, T., SAKATA, M., YOSHIDA, S., KAWAGUCHI, R., KANAYAMA, S., SHIGETOMI, H., HARUTA, S., TSUJI, Y., UEDA, S. & KITANAKA, T. 2008. A randomized study of screening for ovarian

cancer: A multicenter study in Japan. *Obstetrical and Gynaecological Survey*, 63, 635-636.

KOCIOLEK, M., MATERKA, A., STRZELECKI, M. & SZCZYPIŃSKI, P. M. 2001. Discrete Wavelet transform- derived features for digital image texture analysis. *Proc. of International Conference on Signals and Electronic System*, 18-21 September, 163-168.

KOHZUKI, M., KANZAKI, T. & MURATA, Y. 2005. Contrast-enhanced power Doppler sonography of malignant ovarian tumours using harmonic flash-echo imaging: Preliminary experience. *Journal of Clinical Ultrasound*, 33, 237-242.

KOMATSU, T., KONISHI, I., MANDAI, M., TOGASHI, K., KAWAKAMI, S., KONISHI, J. & MORI, T. 1996. Adnexal masses: Transvaginal US and gadolinium-enhanced MR imaging assessment of intratumoural structure. *Radiology*, 198, 109-115.

KREIGER, N., SLOAN, M., COTTERCHIO, M. & PARSONS, P. 1997. Surgical procedures associated with risk of ovarian cancer. *International Journal of Epidemiology*, 26, 710-715.

KUMAR, S. S., MONI, R. S. & RAJEESH, J. Year. Liver tumour diagnosis by gray level and contourlet coefficients texture analysis. In: *Computing, Electronics and Electrical Technologies (ICCEET)*, 2012 International Conference on, 21-22 March 2012. 557-562.

KUMAZAWA, S., UMEZU, T., KANAYAMA, Y., KAMIYAMA, N., SUZUKI, S., MIZUNO, M., KAJIYAMA, H., SHIBATA, K. & KIKKAWA, F. 2012. Contrast-enhanced ultrasonography using Sonazoid® is useful for diagnosis of malignant ovarian tumours: comparison with Doppler ultrasound. *Journal of Medical Ultrasonics*, 1-4.

KURJAK, A. & KUPEŠIĆ, S. 1999. Three dimensional ultrasound and power Doppler in assessment of uterine and ovarian angiogenesis: A prospective study. *Croatian Medical Journal*, 40, 413-420.

KURJAK, A., KUPESIC, S., ANIC, T. & KOSUTA, D. 2000a. Three-dimensional ultrasound and power Doppler improve the diagnosis of ovarian lesions. *Gynaecologic Oncology*, 76, 28-32.

KURJAK, A., KUPESIC, S., SPARAC, V. & KOSUTA, D. 2000b. Three-dimensional ultrasonographic and power Doppler characterization of ovarian lesions. *Ultrasound in Obstetrics and Gynaecology*, 16, 365-371.

KURJAK, A., KUPESIC, S., SPARAC, V., PRKA, M. & BEKAVAC, I. 2003. The detection of stage I ovarian cancer by three-dimensional sonography and power Doppler. *Gynaecologic Oncology*, 90, 258-264.

KURJAK, A. & PREDANIC, M. 1993. New scoring system for prediction of ovarian malignancy based on transvaginal colour Doppler sonography. *Obstetrical and Gynaecological Survey*, 48, 270-272.

- KURJAK, A., PRKA, M., ARENAS, J. M. B., ŠPARAC, V., MERCÉ, L. T., ČORUŠIĆ, A. & IVANČIĆ-KOŠUTA, M. 2005. Three-dimensional ultrasonography and power Doppler in ovarian cancer screening of asymptomatic peri- and postmenopausal women. *Croatian Medical Journal*, 46, 757-764.
- KURODA, H., KAKISAKA, K., KAMIYAMA, N., OIKAWA, T., ONODERA, M., SAWARA, K., OIKAWA, K., ENDO, R., TAKIKAWA, Y. & SUZUKI, K. 2012. Non-invasive determination of hepatic steatosis by acoustic structure quantification from ultrasound echo amplitude. *World Journal of Gastroenterology*, 18, 3889-3895.
- KURTZ, A. B., TSIMIKAS, J. V., TEMPANY, C. M. C., HAMPER, U. M., ARGER, P. H., BREE, R. L., WECHLER, R. J., FRANCIS, I. R., KUHLMAN, J. E., SIEGELMAN, E. S., MITCHELL, D. G., SILVERMAN, S. G., BROWN, D. L., SHETH, S., COLEMAN, B. G., ELLIS, J. H., KURMAN, R. J., CAUDRY, D. J. & MCNELL, B. J. 1999. Diagnosis and staging of ovarian cancer: Comparative values of Doppler and conventional US, CT, and MR imaging correlated with surgery and histopathologic analysis - Report of the radiology diagnostic oncology group. *Radiology*, 212, 19-27.
- LABAN, M., METAWEE, H., ELYAN, A., KAMAL, M., KAMEL, M. & MANSOUR, G. 2007. Three-dimensional ultrasound and three-dimensional power Doppler in the assessment of ovarian tumours. *International Journal of Gynaecology and Obstetrics*, 99, 201-205.
- LABORATORY, N. P. 2005. Measurement Good Practice Guide No. 52, Teddington, Middlesex: National Physical Laboratory.
- LACHIN, J. M. 2004. The role of measurement reliability in clinical trials. *Clinical trials (London, England)*, 1, 553-566.
- LAGALLA, R. & MIDIRI, M. 1998. Image quality control in breast ultrasound. *European Journal of Radiology*, 27, S229-S233.
- LANG, T. & SECIC, M. 2006. How to report statistics in medicine, Philadelphia, American College of Physicians.
- LE, T., GIEDE, C., SALEM, S., LEFEBVRE, G., ROSEN, B., BENTLEY, J., KUPETS, R., POWER, P., RENAUD, M. C., BRYSON, P., DAVIS, D. B., LAU, S., LOTOCKI, R., SENIKAS, V., MORIN, L., BLY, S., BUTT, K., CARGILL, Y. M., DENIS, N., GAGNON, R., HIETALA-COYLE, M. A., LIM, K. I., OUELLET, A. & RACIOT, M. H. 2009. Initial evaluation and referral guidelines for management of pelvic/ovarian masses. *Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstétrique et gynécologie du Canada : JOGC*, 31, 668-680.
- LEA, J. S. & MILLER, D. S. 2001. Optimum screening interventions for Gynaecologic malignancies. *Texas Medicine*, 97, 49-55.
- LEE, T. S., KIM, J. W., PARK, N. H., SONG, Y. S., KANG, S. B. & LEE, H. P. 2005. Assessing clinical performance of Gynaecology residents: Sonographic

evaluation of adnexal masses based on morphological scoring systems. *Ultrasound in Obstetrics and Gynaecology*, 26, 776-779.

LEE, W. L., CHEN, Y. C. & HSIEH, K. S. 2003. Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform. *IEEE Transactions on Medical Imaging*, 22, 382-392.

LEITZMANN, M. F., KOEBNICK, C., DANFORTH, K. N., BRINTON, L. A., MOORE, S. C., HOLLENBECK, A. R., SCHATZKIN, A., LACEY, J. V., JR., LEITZMANN, M. F., KOEBNICK, C., DANFORTH, K. N., BRINTON, L. A., MOORE, S. C., HOLLENBECK, A. R., SCHATZKIN, A. & LACEY, J. V., JR. 2009. Body mass index and risk of ovarian cancer. *Cancer*, 115, 812-22.

LERNER, J. P., TIMOR-TRITSCH, I. E., FEDERMAN, A. & ABRAMOVICH, G. 1994. Transvaginal ultrasonographic characterization of ovarian masses with an improved, weighted scoring system. *American Journal of Obstetrics and Gynaecology*, 170, 81-85.

LERSKI, R. 2006. *Clinical Application of Texture Analysis*. [Accessed 02/10/2015].

LERSKI, R. A., BARNETT, E., MORLEY, P., MILLS, P. R., WATKINSON, G. & MACSWEEN, R. N. M. 1979. Computer analysis of ultrasonic signals in diffuse liver disease. *Ultrasound in Medicine & Biology*, 5, 341-343.

LI, B. & MENG, M. Q. H. 2009. Texture analysis for ulcer detection in capsule endoscopy images. *Image and Vision Computing*, 27, 1336-1342.

LI, P. S., YING, M., CHAN, K. H., CHAN, P. W. & CHU, K. L. 2004. The reproducibility and short-term and long-term repeatability of sonographic measurement of splenic length. *Ultrasound in Medicine and Biology*, 30, 861-866.

LIPSKIE, T. L. 1998. A summary of cancer screening guidelines. *Chronic diseases in Canada*, 19, 112-130.

LIVENS, S., SCHEUNDERS, P., VAN DE WOUWER, G. & VAN DYCK, D. Year. Wavelets for texture analysis, an overview. In, 1997, page 581. 581-585.

LOU, H. Y., MENG, H., ZHU, Q. L., ZHANG, Q. & JIANG, Y. X. 2010. Application values of four risk of malignancy indices in the preoperative evaluation of patients with adnexal masses. *Acta Academiae Medicinae Sinicae*, 32, 297-302.

LOUBEYRE, P., PATEL, S., COPERCINI, M., PETIGNAT, P., DALLENBACH, P. & DUBUISSON, J. B. 2012. Role of sonography in the diagnostic workup of ovarian and adnexal masses except in pregnancy and during ovarian stimulation. *Journal of Clinical Ultrasound*, 40, 424-432.

LU, R. S., TIAN, G. Y., GLEDHILL, D. & WARD, S. 2006. Grinding surface roughness measurement based on the co-occurrence matrix of speckle pattern texture. *Applied Optics*, 45, 8839-8847.

- LUBIN, F., CHETRIT, A., FREEDMAN, L. S., ALFANDARY, E., FISHIER, Y., NITZAN, H., ZULTAN, A. & MODAN, B. 2003. Body mass index at age 18 years and during adult life and ovarian cancer risk. *American Journal of Epidemiology*, 157, 113-120.
- LUO, L. Y., KATSAROS, D., SCORILAS, A., FRACCHIOLI, S., BELLINO, R., VAN GRAMBEREN, M., DE BRUIJN, H., HENRIK, A., STENMAN, U. H., MASSOBRIO, M., VAN DER ZEE, A. G. J., VERGOTE, I. & DIAMANDIS, E. P. 2003. The serum concentration of human kallikrein 10 represents a novel biomarker for ovarian cancer diagnosis and prognosis. *Cancer Research*, 63, 807-811.
- M. HAJEK, M. DEZORTOVA, A. MATERKA & LERSKI, R. 2006. *Texture Analysis for Magnetic Resonance Imaging*. prague: Med4publishing.
- MAJEED, H., RAMZAN, A., IMRAN, F. & MAHFOOZ UR, R. 2011. Validity of resistive index for the diagnosis of malignant ovarian masses. *Journal of the Pakistan Medical Association*, 61, 1104-1107.
- MANCUSO, A., DE VIVO, A., TRIOLO, O. & IRATO, S. 2004. The role of transvaginal ultrasonography and serum CA 125 assay combined with age and hormonal state in the differential diagnosis of pelvic masses. *European Journal of Gynaecological Oncology*, 25, 207-210.
- MANJUNATH, A. P., PRATAPKUMAR, SUJATHA, K. & VANI, R. 2001. Comparison of three risk of malignancy indices in evaluation of pelvic masses. *Gynaecologic Oncology*, 81, 225-229.
- MARRET, H., SAUGET, S., GIRAUDEAU, B., BODY, G. & TRANQUART, F. 2005. Power Doppler vascularity index for predicting malignancy of adnexal masses. *Ultrasound in Obstetrics and Gynaecology*, 25, 508-513.
- MARSHALL, C. 2008. *Diseases and Disorders*, New York, Marshall Cavendish Corporation.
- MATERKA, A. 2004. Texture analysis methodologies for magnetic resonance imaging. *Dialogues in Clinical Neuroscience*, 6, 243-250.
- MATERKA A & M, S. 1998. *Texture Analysis Methods - A Review*. Technical University of Lodz, Institute of Electronics, COST B11 Report, Brussels.
- MATHIAS, J. M., TOFTS, P. S. & LOSSEFF, N. A. 1999. Texture analysis of spinal cord pathology in multiple sclerosis. *Magnetic Resonance in Medicine*, 42, 929-935.
- MCBEE, W. C., ESCOBAR, P. F. & FALCONE, T. 2007. Which ovarian masses need intervention? *Cleveland Clinic Journal of Medicine*, 74, 149-157.
- MCDONALD, J. M., DORAN, S., DESIMONE, C. P., UELAND, F. R., DEPRIEST, P. D., WARE, R. A., SAUNDERS, B. A., PAVLIK, E. J., GOODRICH, S., KRYSCIO, R., J. & NAGELL, J., R. 2010. Predicting Risk of Malignancy in Adnexal Masses. *Obstetrics*, 115, 687-694.

- MEDEIROS, L. R., ROSA, D. D., DA ROSA, M. I. & BOZZETTI, M. C. 2009a. Accuracy of CA 125 in the diagnosis of ovarian tumours: A quantitative systematic review. *European Journal of Obstetrics Gynaecology and Reproductive Biology*, 142, 99-105.
- MEDEIROS, L. R., ROSA, D. D., DA ROSA, M. I., BOZZETTI, M. C., MEDEIROS, L. R., ROSA, D. D., DA ROSA, M. I., X00EA & BOZZETTI, M. C. 2009b. Accuracy of ultrasonography with colour Doppler in ovarian tumour: a systematic quantitative review. *International Journal of Gynaecological Cancer*, 19, 1214-20.
- MEDICAL PHYSICS, W. 2009 (online). Acoustic structure quantification ramps echo resolution [Online]. Available: <http://medicalphysicsweb.org/cws/article/newsfeed/40669> [Accessed 27-02-2014].
- MENON, U., GENTRY-MAHARAJ, A., HALLETT, R., RYAN, A., BURNELL, M., SHARMA, A., LEWIS, S., DAVIES, S., PHILPOTT, S., LOPES, A., GODFREY, K., ORAM, D., HEROD, J., WILLIAMSON, K., SEIF, M. W., SCOTT, I., MOULD, T., WOOLAS, R., MURDOCH, J., DOBBS, S., AMSO, N. N., LEESON, S., CRUICKSHANK, D., MCGUIRE, A., CAMPBELL, S., FALLOWFIELD, L., SINGH, N., DAWNAY, A., SKATES, S. J., PARMAR, M. & JACOBS, I. 2009. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *The Lancet Oncology*, 10, 327-340.
- MENON, U., GENTRY-MAHARAJ, A., RYAN, A., SHARMA, A., BURNELL, M., HALLETT, R., LEWIS, S., LOPEZ, A., GODFREY, K., ORAM, D., HEROD, J., WILLIAMSON, K., SEIF, M., SCOTT, I., MOULD, T., WOOLAS, R., MURDOCH, J., DOBBS, S., AMSO, N., LEESON, S., CRUICKSHANK, D., MCGUIRE, A., CAMPBELL, S., FALLOWFIELD, L., SKATES, S., PARMAR, M. & JACOBS, I. 2008. Recruitment to multicentre trials - Lessons from UKCTOCS: Descriptive study. *BMJ*, 337, 1283-1286.
- MENON, U. & JACOBS, I. J. 2000. Recent developments in ovarian cancer screening. *Current Opinion in Obstetrics and Gynaecology*, 12, 39-42.
- MERZ, E. 1999. Three-dimensional transvaginal ultrasound in Gynaecological diagnosis. *Ultrasound in Obstetrics and Gynaecology*, 14, 81-86.
- MICHAIL, G., KARAHALIOU, A., SKIADOPOULOS, S., KALOGEROPOULOU, C., TERZIS, G., BONIATIS, I., COSTARIDOU, L., KOUROUNIS, G. & PANAYIOTAKIS, G. 2007. Texture analysis of premenopausal and post-menopausal endometrial tissue in grayscale transvaginal ultrasonography. *British Journal of Radiology*, 80, 609-616.
- MILLER, R. W. & UELAND, F. R. 2012. Risk of malignancy in sonographically confirmed ovarian tumours. *Clinical Obstetrics and Gynaecology*, 55, 52-64.

- MIRACLE-MCMAHILL, H. L., CALLE, E. E., KOSINSKI, A. S., RODRIGUEZ, C., WINGO, P. A., THUN, M. J. & HEATH JR, C. W. 1997. Tubal ligation and fatal ovarian cancer in a large prospective cohort study. *American Journal of Epidemiology*, 145, 349-357.
- MODESITT, S. C., PAVLIK, E. J., UELAND, F. R., DEPRIEST, P. D., KRYSCIO, R. J. & VAN NAGELL JR, J. R. 2003. Risk of malignancy in unilocular ovarian cystic tumours less than 10 centimeters in diameter. *Obstetrics and Gynaecology*, 102, 594-599.
- MOHAGHEGH, P. & ROCKALL, A. G. 2012. Imaging strategy for early ovarian cancer: Characterization of adnexal masses with conventional and advanced imaging techniques. *Radiographics*, 32, 1751-1773.
- MOL, B. W. J., BOLL, D., DE KANTER, M., HEINTZ, A. P., SIJMONS, E. A., OEI, S. G., BAL, H. & BRÖLMANN, H. A. M. 2001. Distinguishing the benign and malignant adnexal mass: An external validation of prognostic models. *Gynaecologic Oncology*, 80, 162-167.
- MOOLTHIYA, W., YUENYAO, P., MOOLTHIYA, W. & YUENYAO, P. 2009. The risk of malignancy index (RMI) in diagnosis of ovarian malignancy. *Asian Pacific Journal of Cancer Prevention: Apjcp*, 10, 865-8.
- MOON, W. K., HUANG, Y. S., LO, C. M., HUANG, C. S., BAE, M. S., KIM, W. H., CHEN, J. H. & CHANG, R. F. 2015. Computer-aided diagnosis for distinguishing between triple-negative breast cancer and fibroadenomas based on ultrasound texture features. *Medical Physics*, 42, 3024-3035.
- MOORE, R. G., JABRE-RAUGHLEY, M., BROWN, A. K., ROBISON, K. M., MILLER, M. C., ALLARD, W. J., KURMAN, R. J., BAST, R. C. & SKATES, S. J. 2010. Comparison of a novel multiple marker assay vs the Risk of Malignancy Index for the prediction of epithelial ovarian cancer in patients with a pelvic mass. *American Journal of Obstetrics and Gynaecology*, 203, 228.e1-228.e6.
- MOORE, R. G. & MACLAUGHLAN, S. 2010. Current clinical use of biomarkers for epithelial ovarian cancer. *Current Opinion in Oncology*, 22, 492-497.
- MORAN, A. & MOELLER, H. 2009. Poor progress in cancer control in elderly in UK National cancer Intelligence Network Proceedings of the 2009 NCIN Conference 1-2.
- MØRCH, L. S., LØKKEGAARD, E., ANDREASEN, A., KRÜGER-KJAER, S. & LIDEGAARD, O. 2009. Hormone therapy and ovarian cancer. *JAMA*, 302, 298-305.
- MORGANTE, G., LA MARCA, A., DITTO, A. & DE LEO, V. 1999. Comparison of two malignancy risk indices based on serum CA125, ultrasound score and menopausal status in the diagnosis of ovarian masses. *British Journal of Obstetrics and Gynaecology*, 106, 524-527.

- MORRIS, D. T. 1988. An evaluation of the use of texture measurements for the tissue characterisation of ultrasonic images of in vivo human placentae. *Ultrasound in Medicine and Biology*, 14, 387-395.
- MOSS, E. L., HOLLINGWORTH, J. & REYNOLDS, T. M. 2005. The role of CA125 in clinical practice. *Journal of Clinical Pathology*, 58, 308-312.
- MOTULSKY, H. J. & BROWN, R. E. 2006. Detecting outliers when fitting data with nonlinear regression - A new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics*, 7.
- MOUSAVI, A. S., BORNA, S. & MOEINODDINI, S. 2006. Estimation of probability of malignancy using a logistic model combining colour Doppler ultrasonography, serum CA125 level in women with a pelvic mass. *International Journal of Gynaecological Cancer*, 16, 92-98.
- MYERS ER., BASTAIN LA., HARVRILESKY LJ., KULASINGAM SL., TERPLAN MS., CLINE KE., GARY RN. & DC., M. 2006. Management of Adnexal Mass, Evidence report/Technology Assessment No.130. Rockville: AHRQ Publication.
- MYLES, P. S. & CUI, J. 2007. I. Using the Bland-Altman method to measure agreement with repeated measures. *British Journal of Anaesthesia*, 99, 309-311.
- NAILON, W. H., MCLAUGHLIN, S., SPENCER, T. & RAMO, M. P. Year. Comparative study of textural analysis techniques to characterise tissue from intravascular ultrasound. In, 1996. 303-306.
- NATIONAL CANCER INSTITUTE 2014. A Snapshot of Ovarian Cancer: Incidence and Mortality. In: ONLINE (ed.).
- NELSON, T. R. 2006. Three-dimensional Ultraound imaging. UIA Annual Meeting. Ultrasonic Industry Association.
- NICE, C. G. 2011. CG122 Ovarian cancer: the recognition and initial management of ovarian cancer.
- NOLEN, B. M. & LOKSHIN, A. E. 2013. Biomarker testing for ovarian cancer: Clinical utility of multiplex assays. *Molecular Diagnosis and Therapy*, 17, 139-146.
- O'ROURKE, J. & MAHON, S. M. 2003. A comprehensive look at the early detection of ovarian cancer. *Clinical journal of oncology nursing*, 7, 41-47.
- OKUGAWA, K., HIRAKAWA, T., FUKUSHIMA, K., KAMURA, T., AMADA, S. & NAKANO, H. 2001. Relationship between age, histological type, and size of ovarian tumours. *International Journal of Gynaecology and Obstetrics*, 74, 45-50.
- OLSEN, C. M., GREEN, A. C., WHITEMAN, D. C., SADEGHI, S., KOLAHDOOZ, F. & WEBB, P. M. 2007. Obesity and the risk of epithelial ovarian cancer: A systematic review and meta-analysis. *European Journal of Cancer*, 43, 690-709.

- OLSEN, C. M., NAGLE, C. M., WHITEMAN, D. C., PURDIE, D. M., GREEN, A. C., WEBB, P. M., BOWTELL, D., CHENEVIX-TRENCH, G., GREEN, A., WEBB, P., DEFAZIO, A., GERTIG, D., J., JAYDE, V., BOWES, L., MAMERS, P., SCHMIDT, T., SHIRLEY, H., ROBBIE, M., et al. 2008. Body size and risk of epithelial ovarian and related cancers: A population-based case-control study. *International Journal of Cancer*, 123, 450-456.
- OLSON, S. H., MIGNONE, L., NAKRASEIVE, C., CAPUTO, T. A., BARAKAT, R. R. & HARLAP, S. 2001. Symptoms of ovarian cancer. *Obstetrics and Gynaecology*, 98, 212-217.
- ONODERA, M. 2013. The New Non-Invasive Quantification of Hepatic Steatosis With Morbid Obesity by Acoustic Structure Quantification (ASQ) From Ultrasound Echo Amplitude. *Ultrasound in Medicine & Biology* 39, 5S.
- OOSTERVELD, B. J., THIJSSSEN, J. M., HARTMAN, P. C., ROMIJN, R. L. & ROSENBUSCH, G. J. E. 1991. Ultrasound attenuation and texture analysis of diffuse liver disease: Methods and preliminary results. *Physics in Medicine and Biology*, 36, 1039-1064.
- OPPENHEIM, A. (ed.) 1992. Questionnaire design, interviewing and attitude measurement.: Pinter publishers Ltd,.
- ORTASHI 2008. Cardiff Malignancy Index (CMI)- A novel scoring system for assessing ovarian masses. BSGI Inagural Meeting 2008.
- PALEY, P. J. 2001. Screening for the major malignancies affecting women: Current guidelines. *American Journal of Obstetrics and Gynaecology*, 184, 1021-1030.
- PALMA, M. D., GREGIANIN, M., FIDUCCIA, P., EVANGELISTA, L., CERVINO, A. R., SALADINI, G., BORGATO, L., NICOLETTO, M. O. & ZAGONEL, V. 2012. PET/CT imaging in Gynaecologic malignancies: A critical overview of its clinical impact and our retrospective single centre analysis. *Critical Reviews in Oncology/Haematology*, 83, 84-98.
- PAVLIK, E. J., SAUNDERS, B. A., DORAN, S., MCHUGH, K. W., UELAND, F. R., DESIMONE, C. P., DEPRIEST, P. D., WARE, R. A., KRYSCIO, R. J. & VAN NAGELL JR, J. R. 2009. The search for meaning - Symptoms and transvaginal sonography screening for ovarian cancer predicting malignancy. *Cancer*, 115, 3689-3698.
- PERMUTH-WEY, J. & SELLERS, T. A. 2009. Epidemiology of ovarian cancer. *Methods in Molecular Biology*.
- PETRICOIN III, E. F., ARDEKANI, A. M., HITT, B. A., LEVINE, P. J., FUSARO, V. A., STEINBERG, S. M., MILLS, G. B., SIMONE, C., FISHMAN, D. A., KOHN, E. C. & LIOTTA, L. A. 2002. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*, 359, 572-577.

- PIERSON, R. A. & ADAMS, G. P. 1995. Computer-assisted image analysis, diagnostic ultrasonography and ovulation induction: Strange bedfellows. *Theriogenology*, 43, 105-112.
- PONCELET, C., FAUVET, R., YAZBECK, C., COUTANT, C. & DARAI, E. 2010. Impact of serum tumour marker determination on the management of women with borderline ovarian tumours: Multivariate analysis of a French multicentre study. *European Journal of Surgical Oncology (EJSO)*, 36, 1066-1072.
- POWER, M. & KUYKEN, W. 1998. World Health Organization Quality of Life Assessment (WHOQOL): Development and general psychometric properties. *Social Science and Medicine*, 46, 1569-1585.
- PRAGER, R. W., IJAZ, U. Z., GEE, A. H. & TREECE, G. M. 2010. Three-dimensional ultrasound imaging. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224, 193-223.
- PRYS DAVIES, A., JACOBS, I., WOOLAS, R., FISH, A. & ORAM, D. 1993. The adnexal mass: Benign or malignant? Evaluation of a risk of malignancy index. *British Journal of Obstetrics and Gynaecology*, 100, 927-931.
- RAJENDRA ACHARYA, U., VINITHA SREE, S., KULSHRESHTHA, S., MOLINARI, F., KOH, J. E. W., SABA, L. & SURI, J. S. 2014. GyneScan: An improved online paradigm for screening of ovarian cancer via tissue characterization. *Technology in Cancer Research and Treatment*, 13, 529-540.
- RAMARAJ, M. & RAGHAVAN, S. Year. A survey of wavelet techniques and multiresolution analysis for cancer diagnosis. In: *Computer, Communication and Electrical Technology (ICCCET)*, 2011 International Conference on, 18-19 March 2011 2011. 109-114.
- RAMOS, R. P., NASCIMENTO, M. Z. D. & PEREIRA, D. C. 2012. Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. *Expert Systems with Applications*, 39, 11036-11047.
- RCOG. 2003. Ovarian Cysts in Postmenopausal Women 34. Available: <http://www.rcog.org.uk/womens-health/clinical-guidance/ovarian-cysts-postmenopausal-women-green-top-34> [Accessed 10-02-2011].
- REDMAN, C., DUFFY, S., BROMHAM, N. & FRANCIS, K. 2011. Guidelines: Recognition and initial management of ovarian cancer: Summary of NICE guidance. *BMJ*, 342.
- REZNEK R. 2006. Ultrasound In Ovarian Carcinoma. In: REZNEK R. (ed.) *Cancer of the Ovary*. Cambridge: Cambridge University Press.
- RICCI, P., MARIGLIANO, C., CANTISANI, V., PORFIRI, A., MARCANTONIO, A., LODISE, P., D'AMBROSIO, U., LABBADIA, G., MAGGINI, E., MANCUSO, E., PANZIRONI, G., DI SEGNI, M., FURLAN, C., MASCIANGELO, R. & TALIANI, G. 2013. Ultrasound evaluation of liver fibrosis: Preliminary experience with acoustic structure quantification (ASQ) software. *Valutazione ecografica della*

- fibrosi epatica: esperienza preliminare con acoustic structure quantification (ASQ), 118, 995-1010.
- RIECK, G. C., PUGH, N. D. & FIANDER, A. N. 2006. Power Doppler in the assessment of pelvic masses in a low risk group. *Journal of Obstetrics and Gynaecology*, 26, 222-224.
- RISCH, H. A., MARRETT, L. D. & HOWE, G. R. 1994. Parity, contraception, infertility, and the risk of epithelial ovarian cancer. *American Journal of Epidemiology*, 140, 585-597.
- RODRIGUEZ, C., CALLE, E. E., FAKHRABADI-SHOKOOHI, D., JACOBS, E. J. & THUN, M. J. 2002. Body mass index, height, and the risk of ovarian cancer mortality in a prospective cohort of postmenopausal women. *Cancer Epidemiology Biomarkers and Prevention*, 11, 822-828.
- ROELLINGER JR, F. X., KAHVECI, A. E., CHANG, J. K., HARLOW, C. A., DWYER III, S. J. & LODWICK, G. S. 1973. Computer analysis of chest radiographs. *Computer Graphics and Image Processing*, 2, 232-251.
- RONCO, D. A., MANAHAN, K. J. & GEISLER, J. P. 2011. Ovarian cancer risk assessment: a tool for preoperative assessment. *European Journal of Obstetrics & Gynaecology and Reproductive Biology*, 158, 325-329.
- ROSEN, D. G., WANG, L., ATKINSON, J. N., YU, Y., LU, K. H., DIAMANDIS, E. P., HELLSTROM, I., MOK, S. C., LIU, J. & BAST JR, R. C. 2005. Potential markers that complement expression of CA125 in epithelial ovarian cancer. *Gynaecologic Oncology*, 99, 267-277.
- ROSENBLATT, K. A. & THOMAS, D. B. 1996. Reduced risk of ovarian cancer in women with a tubal ligation or hysterectomy. *Cancer Epidemiology Biomarkers and Prevention*, 5, 933-935.
- ROSSI, A., BRAGHIN, C., SOLDANO, F., ISOLA, M., CAPODICASA, V., LONDERO, A. P., FORZANO, L. & MARCHESONI, D. 2011. A proposal for a new scoring system to evaluate pelvic masses: Pelvic Masses Score (PMS). *European Journal of Obstetrics & Gynaecology and Reproductive Biology*, 157, 84-88.
- RUBIN, J. M. 1994. Spectral Doppler US. Radiographics : a review publication of the Radiological Society of North America, Inc, 14, 139-150.
- RUFFORD, B. & JACOBS, I. 2003. Ovarian cysts in postmenopausal women. In: *GYNAECOLOGISTS*, R. C. O. O. A. (ed.).
- RUMACK, C. (ed.) 2005. *Diagnostic Ultrasound*: Mosby inc.
- RUMACK, C. 2011. *Diagnostic Ultrasound*, Mosby inc.
- RUSSELL, P. (ed.) 1994. *Surface Epithelial-Stromal Tumours of the Ovary*, New York: Springer-Verilog.

- SAITO, M., AOKI, D., SUSUMU, N., SUZUKI, A., SUZUKI, N., UDAGAWA, Y. & NOZAWA, S. 2005. Efficient screening for ovarian cancers using a combination of tumour markers CA602 and CA546. *International Journal of Gynaecological Cancer*, 15, 37-44.
- SANDERS, R. 1998. *Clinical Sonography (a practical guide)*, New York, Lippincott, Philadelphia.
- SARKAR, M. & WOLF, M. G. 2012. Simple ovarian cysts in postmenopausal women: Scope of conservative management. *European Journal of Obstetrics Gynaecology and Reproductive Biology*, 162, 75-78.
- SASSONE, A. M., TIMOR-TRITSCH, I. E., ARTNER, A., WESTHOFF, C. & WARREN, W. B. 1991. Transvaginal sonographic characterization of ovarian disease: Evaluation of a new scoring system to predict ovarian malignancy. *Obstetrics and Gynaecology*, 78, 70-76.
- SAUNDAR, S., SNELL, K., DEEKS, J., C., D., A., S., TIMMERMAN, D. & MENON, U. 2015. ROCKETTS Study. Birmingham: University of Birmingham Clinical Trial Units (BCTU).
- SAUNDERS, B. A., PODZIELINSKI, I., WARE, R. A., GOODRICH, S., DESIMONE, C. P., UELAND, F. R., SEAMON, L., UBELLACKER, J., PAVLIK, E. J., KRYSZCIO, R. J. & VAN NAGELL JR, J. R. 2010. Risk of malignancy in sonographically confirmed septated cystic ovarian tumours. *Gynaecologic Oncology*, 118, 278-282.
- SCHOUTEN, L. J., RIVERA, C., HUNTER, D. J., SPIEGELMAN, D., ADAMI, H. O., ARSLAN, A., BEESON, W. L., VAN DEN BRANDT, P. A., BURING, J. E., FOLSOM, A. R., FRASER, G. E., FREUDENHEIM, J. L., GOL%OHM, R. A., HANKINSON, S. E., LACEY JR, J. V., LEITZMANN, M., LUKANOVA, A., MARSHALL, J. R., MILLER, A. B., PATEL, A. V., RODRIGUEZ, C., ROHAN, T. E., ROSS, J. A., WOLK, A., ZHANG, S. M. & SMITH-WARNER, S. A. 2008. Height, body mass index, and ovarian cancer: A pooled analysis of 12 cohort studies. *Cancer Epidemiology Biomarkers and Prevention*, 17, 902-912.
- SCHRECENGOST, A. 2002. Ovarian Mass—Benign or Malignant? *AORN*, 76, 789-806.
- SCHUTTER, E. M. J., SOHN, C., KRISTEN, P., MÖBUS, V., CROMBACH, G., KAUFMANN, M., CAFFIER, H., KREIENBERG, R., VERSTRAETEN, A. A. & KENEMANS, P. 1998. Estimation of probability of malignancy using a logistic model combining physical examination, ultrasound, serum CA 125, and serum CA 72-4 in postmenopausal women with a pelvic mass: An international multicentre study. *Gynaecologic Oncology*, 69, 56-63.
- SCHWARTZ, P. E., CHAMBERS, J. T., TAYLOR, K. J., PELLERITO, J., HAMMERS, L., COLE, L. A., YANG-FENG, T. L., SMITH, P., MAYNE, S. T. & MAKUCH, R. 1991. Early detection of ovarian cancer: Preliminary results of the Yale Early Detection Program. *Yale Journal of Biology and Medicine*, 64, 573-582.

SCONFIENZA, L. M., PERRONE, N., DELNEVO, A., LACELLI, F., MUROLO, C., GANDOLFO, N. & SERAFINI, G. 2010. Diagnostic value of contrast-enhanced ultrasonography in the characterization of ovarian tumours. *Journal of Ultrasound*, 13, 9-15.

SELECT STATISTICAL, S. 2015. The Importance and Effect of Sample Size [Online]. Exeter. Available: <http://www.select-statistics.co.uk/article/blog-post/the-importance-and-effect-of-sample-size> [Accessed 22-10-2015].

SHARMA, A., GENTRY-MAHARAJ, A., BURNELL, M., FOURKALA, E.-O., CAMPBELL, S., AMSO, N. N., SEIF, M. W., RYAN, A., PARMAR, M., JACOBS, I. J. & MENON, U. Published Online 2011. Assessing the malignant potential of the ovarian inclusion cysts in postmenopausal women within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a prospective cohort study. *Gynaecologic Oncology* [Online].

SHARMA, A. & MENON, U. 2006. Screening for gynaecological cancers. *European Journal of Surgical Oncology (EJSO)*, 32, 818-824.

SHARMA, M. & SINGH, S. 2001. Evaluation of Texture Methods for Image Analysis. Seventh Australian and New Zealand Intelligent Information Systems Conference 18-21 November.

SHEN, S., DENARDO, G. L., DENARDO, S. J., YUAN, A., DENARDO, D. A. & LAMBORN, K. R. 1997. Reproducibility of operator processing for radiation dosimetry. *Nuclear Medicine and Biology*, 24, 77-83.

SHUNG, K. K. 2006. *Diagnostic Ultrasound: Imaging and Blood Flow Measurements*, London, Taylor & Francis.

SILVERMAN, W. A. 1985. *Human experimentation: a guided step into the unknown*, Oxford, University press.

SINGH, J., PIERSON, R. A. & ADAMS, G. P. 1997. Ultrasound image attributes of the bovine corpus luteum: Structural and functional correlates. *Journal of Reproduction and Fertility*, 109, 35-44.

SINGH, M. & SINGH, S. 2002. Spatial texture analysis: A comparative study. *Proceedings - International Conference on Pattern Recognition*, 16, 676-679.

SINHA A, DREW F, LIM K & N., P. 2015. Retrospective analysis of suspicious pelvic masses using the pelvic mass index (PMI) scoring system: the Cardiff experience- 2007 to 2014. Cardiff.

SIRINIVASAN, G. N. & SHOBHA, G. 2008. Statistical Texture Analysis. *Proceeding of world academy of Science, Engineering and Technology* 36, 1264-1269.

SKATES, S. J., HORICK, N., YU, Y., XU, F. J., BERCHUCK, A., HAVRILESKY, L. J., DE BRUIJN, H. W. A., VAN DER ZEE, A. G. J., WOOLAS, R. P., JACOBS, I. J., ZHANG, Z. & BAST JR, R. C. 2004. Preoperative sensitivity and specificity for

early-stage ovarian cancer when combining cancer antigen CA-125II, CA 15-3, CA 72-4, and macrophage colony-stimulating factor using mixtures of multivariate normal distributions. *Journal of Clinical Oncology* 22, 4059-4066.

SKATES, S. J., MENON, U., MACDONALD, N., ROSENTHAL, A. N., ORAM, D. H., KNAPP, R. C. & JACOBS, I. J. 2003. Calculation of the risk of ovarian cancer from serial CA-125 values for preclinical detection in postmenopausal women. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 21, 206s-210s.

SMITH, L. H., MORRIS, C. R., YASMEEN, S., PARIKH-PATEL, A., CRESS, R. D. & ROMANO, P. S. 2005. Ovarian cancer: Can we make the clinical diagnosis earlier? *Cancer*, 104, 1398-1407.

SMUTEK, D., ŠÁRA, R., SUCHARDA, P., TJAHJADI, T. & ŠVEC, M. 2003. Image texture analysis of sonograms in chronic inflammations of thyroid gland. *Ultrasound in Medicine & Biology*, 29, 1531-1543.

SOHAIB, S. A., MILLS, T. D., SAHDEV, A., WEBB, J. A. W., VANTRAPPEN, P. O., JACOBS, I. J. & REZNEK, R. H. 2005. The role of magnetic resonance imaging and ultrasound in patients with adnexal masses. *Clinical Radiology*, 60, 340-348.

SOKALSKA, A., TIMMERMAN, D., TESTA, A. C., VAN HOLSBEKE, C., LISSONI, A. A., LEONE, F. P. G., JURKOVIC, D. & VALENTIN, L. 2009. Diagnostic accuracy of transvaginal ultrasound examination for assigning a specific diagnosis to adnexal masses. *Ultrasound in Obstetrics and Gynaecology*, 34, 462-470.

STANY, M. P., LARRY MAXWELL, G. & ROSE, G. S. 2010. Clinical decision making using ovarian cancer risk assessment. *American Journal of Roentgenology*, 194, 337-342.

SUDARSHAN, V. K., MOOKIAH, M. R. K., ACHARYA, U. R., CHANDRAN, V., MOLINARI, F., FUJITA, H. & NG, K. H. 2016. Application of wavelet techniques for cancer diagnosis using ultrasound images: A Review. *Computers in Biology and Medicine*, 69, 97-111.

SWINSCOW, T. D. V. 1997. *Statistics at Square One*, BMJ Publishing Group.

SZCZYPIŃSKI, P. M., STRZELECKI, M., MATERKA, A. & KLEPACZKO, A. 2009. MaZda—A software package for image texture analysis. *Computer Methods and Programs in Biomedicine*, 94, 66-76.

TATE, A. R., NICHOLSON, A. & CASSELL, J. A. 2010. Are GPs under-investigating older patients presenting with symptoms of ovarian cancer? Observational study using General Practice Research Database. *British Journal of Cancer*, 102, 947-951.

TEMPE, A., SINGH, S., WADHWA, L. & GARG, A. 2006. Conventional and colour Doppler sonography in preoperative assessment of ovarian tumours. *International Journal of Gynaecology and Obstetrics*, 92, 64-68.

- TESAŘ, L., SHIMIZU, A., SMUTEK, D., KOBATAKE, H. & NAWANO, S. 2008. Medical image analysis of 3D CT images based on extension of Haralick texture features. *Computerized Medical Imaging and Graphics*, 32, 513-520.
- TESTA, A. C., AJOSSA, S., FERRANDINA, G., FRUSCELLA, E., LUDOVISI, M., MALAGGESE, M., SCAMBIA, G., MELIS, G. B. & GUERRIERO, S. 2005. Does quantitative analysis of three-dimensional power Doppler angiography have a role in the diagnosis of malignant pelvic solid tumours? A preliminary study. *Ultrasound in Obstetrics and Gynaecology*, 26, 67-72.
- TESTA, A. C., TIMMERMAN, D., VAN BELLE, V., FRUSCELLA, E., VAN HOLSBEKE, C., SAVELLI, L., FERRAZZI, E., LEONE, F. P. G., MARRET, H., TRANQUART, F., EXACOUSTOS, C., NAZZARO, G., BOKOR, D., MAGRI, F., VAN HUFFEL, S., FERRANDINA, G. & VALENTIN, L. 2009. Intravenous contrast ultrasound examination using contrast-tuned imaging (CnTI™) and the contrast medium SonoVue® for discrimination between benign and malignant adnexal masses with solid components. *Ultrasound in Obstetrics and Gynaecology*, 34, 699-710.
- THEOCHARAKIS, P., GLOTSOS, D., KALATZIS, I., KOSTOPOULOS, S., GEORGIADIS, P., SIFAKI, K., TSAKOURIDOU, K., MALAMAS, M., DELIBASIS, G., CAVOURAS, D. & NIKIFORIDIS, G. 2009. Pattern recognition system for the discrimination of multiple sclerosis from cerebral microangiopathy lesions based on texture analysis of magnetic resonance images. *Magnetic Resonance Imaging*, 27, 417-422.
- TIFFEN, J. & MAHON, S. M. 2005. Ovarian cancer screening: are there any options? *Clinical journal of oncology nursing*, 9, 369-372.
- TIMMERMAN, D. 2000. Lack of standardization in Gynaecological ultrasonography. *Ultrasound in Obstetrics and Gynaecology*, 16, 395-398.
- TIMMERMAN, D. 2004. The use of mathematical models to evaluate pelvic masses; can they beat an expert operator? *Best Practice & Research Clinical Obstetrics & Gynaecology*, 18, 91-104.
- TIMMERMAN, D., TESTA, A. C., BOURNE, T., AMEYE, L., JURKOVIC, D., VAN HOLSBEKE, C., PALADINI, D., VAN CALSTER, B., VERGOTE, I., VAN HUFFEL, S. & VALENTIN, L. 2008. Simple ultrasound-based rules for the diagnosis of ovarian cancer. *Ultrasound in Obstetrics and Gynaecology*, 31, 681-690.
- TIMMERMAN, D., VAN CALSTER, B., JURKOVIC, D., VALENTIN, L., TESTA, A. C., BERNARD, J. P., VAN HOLSBEKE, C., VAN HUFFEL, S., VERGOTE, I. & BOURNE, T. 2007. Inclusion of CA-125 does not improve mathematical models developed to distinguish between benign and malignant adnexal tumours. *Journal of Clinical Oncology*, 25, 4194-4200.
- TIMMERMAN, D., VAN CALSTER, B., TESTA, A. C., GUERRIERO, S., FISCHEROVA, D., LISSONI, A. A., VAN HOLSBEKE, C., FRUSCIO, R., CZEKIERDOWSKI, A., JURKOVIC, D., SAVELLI, L., VERGOTE, I., BOURNE, T., VAN HUFFEL, S. & VALENTIN, L. 2010. Ovarian cancer prediction in adnexal

masses using ultrasound-based logistic regression models: A temporal and external validation study by the IOTA group. *Ultrasound in Obstetrics and Gynaecology*, 36, 226-234.

TIMMERMAN, D., VAN CALSTER, B., VERGOTE, I., VAN HOORDE, K., VAN GORP, T., VALENTIN, L. & BOURNE, T. 2011. Performance of the American college of obstetricians and gynaecologists ovarian tumour referral guidelines with a multivariate index assay. *Obstetrics and Gynaecology*, 118, 1179-1181.

TINGULSTAD, S., HAGEN, B., SKJELDESTAD, F. E., HALVORSEN, T., NUSTAD, K. & ONSRUD, M. 1999. The risk-of-malignancy index to evaluate potential ovarian cancers in local hospitals. *Obstetrics and Gynaecology*, 93, 448-452.

TINGULSTAD, S., HAGEN, B., SKJELDESTAD, F. E., ONSRUD, M., KISERUD, T., HALVORSEN, T. & NUSTAD, K. 1996. Evaluation of a risk of malignancy index based on serum CA125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses. *British Journal of Obstetrics and Gynaecology*, 103, 826-831.

TOGASHI, K. 2003. Ovarian cancer: The clinical role of US, CT, and MRI. *European Radiology*, 13, L87-L104.

TOLE, N. 2005. Ultrasound Beam Shape. In: OSTENSEN, H. (ed.) *Basic Physics of ultrasonographic imaging*

Reston, USA: World Health Organization, the International Society of Radiographers and Radiological Technologists.

TORIWAKI, J.-I., SUENAGA, Y., NEGORO, T. & FUKUMURA, T. 1973. PATTERN RECOGNITION OF CHEST X-RAY IMAGES. 125-137.

TOURASSI, G. D. 1999. Journey toward computer-aided diagnosis: Role of image texture analysis. *Radiology*, 213, 317-320.

TOYODA, H., KUMADA, T., KAMIYAMA, N., SHIRAKI, K., TAKASE, K., YAMAGUCHI, T. & HACHIYA, H. 2009. B-mode ultrasound with algorithm based on statistical analysis of signals: Evaluation of liver fibrosis in patients with chronic hepatitis C. *American Journal of Roentgenology*, 193, 1037-1043.

TRIVERS, K. F., STEWART, S. L., PEIPINS, L., RIM, S. H. & WHITE, M. C. 2009. Expanding the public health research agenda for ovarian cancer. *Journal of Women's Health*, 18, 1299-1305.

TSAI, D. Y. & KOJIMA, K. 2005. Measurements of texture features of medical images and its application to computer-aided diagnosis in cardiomyopathy. *Measurement: Journal of the International Measurement Confederation*, 37, 284-292.

TUCERYAN, M. & JAIN, A. K. 1998. *The Handbook of Pattern Recognition and Computer Vision*, World Scientific Publishing Co.

TWICKLER, D. M., FORTE, T. B., SANTOS-RAMOS, R., MCINTIRE, D., HARRIS, P. & MILLER, D. S. 1999. The ovarian tumour index predicts risk for malignancy. *Cancer*, 86, 2280-2290.

TWICKLER, D. M. & MOSCHOS, E. 2010. Ultrasound and assessment of ovarian cancer risk. *American Journal of Roentgenology*, 194, 322-329.

UELAND, F. R., DEPRIEST, P. D., PAVLIK, E. J., KRYSCIO, R. J. & VAN NAGELL JR, J. R. 2003. Preoperative differentiation of malignant from benign ovarian tumours: The efficacy of morphology indexing and Doppler flow sonography. *Gynaecologic Oncology*, 91, 46-50.

ULUSOY, S., AKBAYIR, O., NUMANOGLU, C., ULUSOY, N., ODABAS, E., GULKILIK, A., ULUSOY, S., AKBAYIR, O., NUMANOGLU, C., ULUSOY, N., ODABAS, E. & GULKILIK, A. 2007. The risk of malignancy index in discrimination of adnexal masses. *International Journal of Gynaecology & Obstetrics*, 96, 186-91.

URBAN, N. 1999. Screening for ovarian cancer. *British Medical Journal*, 319, 1317-1318.

URBAN, N., THORPE, J. D., BERGAN, L. A., FORREST, R. M., KAMPANI, A. V., SCHOLLER, N., O'BRIANT, K. C., ANDERSON, G. L., CRAMER, D. W., BERG, C. D., MCINTOSH, M. W., HARTGE, P. & DRESCHER, C. W. 2011. Potential role of HE4 in multimodal screening for epithelial ovarian cancer. *Journal of the National Cancer Institute*, 103, 1630-1634.

VAES, E., MANCHANDA, R., AUTIER, P., NIR, R., NIR, D., BLEIBERG, H., ROBERT, A. & MENON, U. 2012. Differential diagnosis of adnexal masses: Sequential use of the risk of malignancy index and HistoScanning, a novel computer-aided diagnostic tool. *Ultrasound in Obstetrics and Gynaecology*, 39, 91-98.

VALENTIN, L. 2000. Comparison of Lerner score, Doppler ultrasound examination, and their combination for discrimination between benign and malignant adnexal masses. *Ultrasound in Obstetrics and Gynaecology*, 15, 143-147.

VALENTIN, L. 2004. Use of morphology to characterize and manage common adnexal masses. *Best Practice and Research: Clinical Obstetrics and Gynaecology*, 18, 71-89.

VALENTIN, L., AMEYE, L., JURKOVIC, D., METZGER, U., X00E, CURU, F., VAN HUFFEL, S., TIMMERMAN, D., VALENTIN, L., AMEYE, L., JURKOVIC, D., METZGER, U., X00E, CURU, F., VAN HUFFEL, S. & TIMMERMAN, D. 2006. Which extrauterine pelvic masses are difficult to correctly classify as benign or malignant on the basis of ultrasound findings and is there a way of making a correct diagnosis? *Ultrasound in Obstetrics & Gynaecology*, 27, 438-44.

VALENTIN, L., AMEYE, L., SAVELLI, L., FRUSCIO, R., LEONE, F. P. G., CZEKIERDOWSKI, A., LISSONI, A. A., FISCHEROVA, D., GUERRIERO, S., VAN HOLSBEKE, C., VAN HUFFEL, S. & TIMMERMAN, D. 2011. Adnexal masses difficult to classify as benign or malignant using subjective assessment of

grey-scale and Doppler ultrasound findings: Logistic regression models do not help. *Ultrasound in Obstetrics and Gynaecology*, 38, 456-465.

VALENTIN, L., AMEYE, L., SAVELLI, L., FRUSCIO, R., LEONE, F. P. G., CZEKIERDOWSKI, A., LISSONI, A. A., FISCHEROVA, D., GUERRIERO, S., VAN HOLSBEKE, C., VAN HUFFEL, S. & TIMMERMAN, D. 2013. Unilocular adnexal cysts with papillary projections but no other solid components: Is there a diagnostic method that can classify them reliably as benign or malignant before surgery? *Ultrasound in Obstetrics and Gynaecology*, 41, 570-581.

VALENTIN, L., SLADKEVICIUS, P. & MARSAL, K. 1994. Limited contribution of Doppler velocimetry to the differential diagnosis of extrauterine pelvic tumours. *Obstetrics and Gynaecology*, 83, 425-433.

VAN CALSTER, B., TIMMERMAN, D., BOURNE, T., TESTA, A. C., VAN HOLSBEKE, C., DOMALI, E., JURKOVIC, D., NEVEN, P., VAN HUFFEL, S. & VALENTIN, L. 2007. Discrimination between benign and malignant adnexal masses by specialist ultrasound examination versus serum CA-125. *Journal of the National Cancer Institute*, 99, 1706-1714.

VAN CALSTER, B., VAN HOORDE, K., VALENTIN, L., TESTA, A. C., FISCHEROVA, D., VAN HOLSBEKE, C., SAVELLI, L., FRANCHI, D., EPSTEIN, E., KAIJSER, J., VAN BELLE, V., CZEKIERDOWSKI, A., GUERRIERO, S., FRUSCIO, R., LANZANI, C., SCALA, F., BOURNE, T. & TIMMERMAN, D. 2014a. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: Prospective multicentre diagnostic study. *BMJ (Online)*, 349.

VAN CALSTER, B., VAN HOORDE, K., VALENTIN, L., TESTA, A. C., FISCHEROVA, D., VAN HOLSBEKE, C., SAVELLI, L., FRANCHI, D., EPSTEIN, E., KAIJSER, J., VAN BELLE, V., CZEKIERDOWSKI, A., GUERRIERO, S., FRUSCIO, R., LANZANI, C., SCALA, F., BOURNE, T., TIMMERMAN, D. & INTERNATIONAL OVARIAN TUMOUR ANALYSIS, G. 2014b. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. *BMJ (Clinical research ed.)*, 349.

VAN DEN AKKER, P. A. J., AALDERS, A. L., SNIJDERS, M. P. L. M., KLUIVERS, K. B., SAMLAL, R. A. K., VOLLEBERGH, J. H. A. & MASSUGER, L. F. A. G. 2010. Evaluation of the risk of malignancy index in daily clinical management of adnexal masses. *Gynaecologic Oncology*, 116, 384-388.

VAN GORP, T., TIMMERMAN, D. & VERGOTE, I. 2011. Reply: The performance of the risk of ovarian malignancy algorithm. *British Journal of Cancer*, 105, 187-188.

VAN GORP, T., VELDMAN, J., VAN CALSTER, B., CADRON, I., LEUNEN, K., AMANT, F., TIMMERMAN, D. & VERGOTE, I. 2012. Subjective assessment by ultrasound is superior to the risk of malignancy index (RMI) or the risk of ovarian

malignancy algorithm (ROMA) in discriminating benign from malignant adnexal masses. *European Journal of Cancer*, 48, 1649-1656.

VAN HOLSBEKE, C., DAEMEN, A., YAZBEK, J., HOLLAND, T. K., BOURNE, T., MESENS, T., LANNOO, L., BOES, A. S., JOOS, A., VAN DE VIJVER, A., ROGGEN, N., DE MOOR, B., DE JONGE, E., TESTA, A. C., VALENTIN, L., JURKOVIC, D. & TIMMERMAN, D. 2010. Ultrasound experience substantially impacts on diagnostic performance and confidence when adnexal masses are classified using pattern recognition. *Gynaecologic and Obstetric Investigation*, 69, 160-168.

VAN HOLSBEKE, C., DAEMEN, A., YAZBEK, J., HOLLAND, T. K., BOURNE, T., MESENS, T., LANNOO, L., DE MOOR, B., DE JONGE, E., TESTA, A. C., VALENTIN, L., JURKOVIC, D. & TIMMERMAN, D. 2009. Ultrasound methods to distinguish between malignant and benign adnexal masses in the hands of examiners with different levels of experience. *Ultrasound in Obstetrics and Gynaecology*, 34, 454-461.

VAN NAGELL JR, J. R. & DEPRIEST, P. D. 2005. Management of adnexal masses in postmenopausal women. *American Journal of Obstetrics and Gynaecology*, 193, 30-35.

VAN NAGELL JR, J. R., DEPRIEST, P. D., UELAND, F. R., DESIMONE, C. P., COOPER, A. L., MCDONALD, J. M., PAVLIK, E. J. & KRYSCIO, R. J. 2007. Ovarian cancer screening with annual transvaginal sonography: Findings of 25,000 women screened. *Cancer*, 109, 1887-1896.

VAN NAGELL JR, J. R. & UELAND, F. R. 1999. Ultrasound evaluation of pelvic masses: Predictors of malignancy for the general gynaecologist. *Current Opinion in Obstetrics and Gynaecology*, 11, 45-49.

VAN TRAPPEN, P. O., RUFFORD, B. D., MILLS, T. D., SOHAIB, S. A., WEBB, J. A. W., SAHDEV, A., CARROLL, M. J., BRITTON, K. E., REZNEK, R. H. & JACOBS, I. J. 2007. Differential diagnosis of adnexal masses: Risk of malignancy index, ultrasonography, magnetic resonance imaging, and radioimmunosciintigraphy. *International Journal of Gynaecological Cancer*, 17, 61-67.

VAR, T., TONGUC, E. A., UGUR, M., ALTINBAS, S. & TOKMAK, A. 2012. Tumour markers panel and tumour size of ovarian dermoid tumours in reproductive age. *Bratislavské lekárske listy*, 113, 95-98.

VARRAS, M. 2004. Benefits and limitations of ultrasonographic evaluation of uterine adnexal lesions in early detection of ovarian cancer. *Clinical & Experimental Obstetrics & Gynaecology*, 31, 85-98.

VERGOTE, I., AMANT, F., AMEYE, L., TIMMERMAN, D., VERGOTE, I., AMANT, F., X00E, X00E, RIC, AMEYE, L. & TIMMERMAN, D. 2009. Screening for ovarian carcinoma: not quite there yet. *Lancet Oncology*, 10, 308-9.

- VICAS, C., LUPSOR, M., SOCACIU, M., BADEA, R. & NEDEVSCI, S. Year. Texture analysis as a noninvasive tool for fibrosis assessment in chronic hepatitis C. influence of expert dependent variability over the performance of texture analysis. In: Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on, 25-27 Aug. 2011 2011. 205-212.
- VIDYA, K. S., NG, E. Y. K., ACHARYA, U. R., CHOU, S. M., TAN, R. S. & GHISTA, D. N. 2015. Computer-aided diagnosis of Myocardial Infarction using ultrasound images with DWT, GLCM and HOS methods: A comparative study. *Computers in Biology and Medicine*, 62, 86-93.
- VINCE, D. G., DIXON, K. J., COTHREN, R. M. & CORNHILL, J. F. 2000. Comparison of texture analysis methods for the characterization of coronary plaques in intravascular ultrasound images. *Computerized Medical Imaging and Graphics*, 24, 221-229.
- VINE, M. F., CALINGAERT, B., BERCHUCK, A. & SCHILDKRAUT, J. M. 2003. Characterization of prediagnostic symptoms among primary epithelial ovarian cancer cases and controls. *Gynaecologic Oncology*, 90, 75-82.
- VISINTIN, I., FENG, Z., LONGTON, G., WARD, D. C., ALVERO, A. B., LAI, Y., TENTHOREY, J., LEISER, A., FLORES-SAAIB, R., YU, H., AZORI, M., RUTHERFORD, T., SCHWARTZ, P. E. & MOR, G. 2008. Diagnostic markers for early detection of ovarian cancer
Clinical Cancer Research, 14, 1065-1072.
- WAKAHARA, F., KIKKAWA, F., NAWA, A., TAMAKOSHI, K., INO, K., MAEDA, O., KAWAI, M. & MIZUTANI, S. 2001. Diagnostic efficacy of tumour markers, sonography, and intraoperative frozen section for ovarian tumours. *Gynaecologic and Obstetric Investigation*, 52, 147-152.
- WANG, L. M., SONG, H., SONG, X. & ZHOU, X. B. 2012. An improved risk of malignancy index in diagnosis of adnexal mass. *Chinese Medical Journal*, 125, 533-535.
- WANG, Y., TANIGUCHI, N. & ITOH, K. 2002. The development and clinical application of a new texture analysis system for acoustic tissue characterization. *Ultrasound International*, 8, 81-105.
- WANG, Y. Z., WANG, X. M., LI, Y. Y. & OU, G. C. 2013. Diagnostic value of acoustic structure quantification technology in diffuse liver diseases. *World Chinese Journal of Digestology*, 21, 448-453.
- WEBB, J. 2007. *Ultrasound in Ovarian Carcinoma*. Cambridge: Cambridge university press.
- WHITTINGHAM, T. & MARTIN, K. 2010. *Transducer and beam forming*, Cambridge, Cambridge University Press.

- WILSON, W. D., VALET, A. S., ANDREOTTI, R. F., GREEN-JARVIS, B., LYSCHCHIK, A. & FLEISCHER, A. C. 2006. Sonographic quantification of ovarian tumour vascularity. *Journal of Ultrasound in Medicine*, 25, 1577-1581.
- WOODWARD, P. J., HOSSEINZADEH, K. & SAENGER, J. S. 2004. Radiologic Staging of Ovarian Carcinoma with Pathologic Correlation. *Radiographics*, 24, 225-246.
- WOOLAS, R. P., ORAM, D. H., JEYARAJAH, A. R., BAST JR, R. C. & JACOBS, I. J. 1999. Ovarian cancer identified through screening with serum markers but not by pelvic imaging. *International Journal of Gynaecological Cancer*, 9, 497-501.
- WU, H.-K., TSENG, H.-S., CHEN, L.-S., KUO, S.-J., YEN, P.-L., HUANG, Y.-L. & CHEN, D.-R. 2014. Flow Index of 3D Breast Ultrasound was Strongly Correlated with Axillary Lymph Node Metastasis in the Absence of Lymphovascular Invasion. *Journal of Medical Imaging and Health Informatics*, 4, 66-70.
- WU, M. H., CHENG, Y. C., CHANG, C. H., KO, H. C. & CHANG, F. M. 2012. Three-dimensional Ultrasound in Evaluation of the Ovary. *Journal of Medical Ultrasound*, 20, 136-141.
- XIAN, G.-M. 2010. An identification method of malignant and benign liver tumours from ultrasonography based on GLCM texture features and fuzzy SVM. *Expert Systems with Applications*, 37, 6737-6741.
- XU, Y., VAN BEEK, E. J. R., HWANJO, Y., GUO, J., MCLENNAN, G. & HOFFMAN, E. A. 2006. Computer-aided Classification of Interstitial Lung Diseases Via MDCT: 3D Adaptive Multiple Feature Method (3D AMFM). *Academic Radiology*, 13, 969-978.
- YAMAMOTO, Y., YAMADA, R., OGURI, H., MAEDA, N. & FUKAYA, T. 2009. Comparison of four malignancy risk indices in the preoperative evaluation of patients with pelvic masses. *European Journal of Obstetrics Gynaecology and Reproductive Biology*, 144, 163-167.
- YAMASHITA, Y., HATANAKA, Y., TORASHIMA, M., TAKAHASHI, M., MIYAZAKI, K. & OKAMURA, H. 1997. Characterization of sonographically indeterminate ovarian tumours with MR imaging: A logistic regression analysis. *Acta Radiologica*, 38, 572-577.
- YAMASHITA, Y., TORASHIMA, M., HATANAKA, Y., HARADA, M., HIGASHIDA, Y., TAKAHASHI, M., MIZUTANI, H., TASHIRO, H., IWAMASA, J., MIYAZAKI, K. & OKAMURA, H. 1995. Adnexal masses: Accuracy of characterization with transvaginal US and precontrast and postcontrast MR imaging. *Radiology*, 194, 557-565.
- YAZBEK, J., AMEYE, L., TIMMERMAN, D., TESTA, A. C., VALENTIN, L., HOLLAND, T. K., VAN HOLSBEKE, C. & JURKOVIC, D. 2010. Use of ultrasound pattern recognition by expert operators to identify borderline ovarian tumours: A

study of diagnostic performance and interobserver agreement. *Ultrasound in Obstetrics and Gynaecology*, 35, 84-88.

YAZBEK, J., ASLAM, N., TAILOR, A., HILLABY, K., RAJU, K. S. & JURKOVIC, D. 2006. A comparative study of the risk of malignancy index and the ovarian crescent sign for the diagnosis of invasive ovarian cancer. *Ultrasound in Obstetrics and Gynaecology*, 28, 320-324.

YAZBEK, J., HELMY, S., BEN-NAGI, J., HOLLAND, T., SAWYER, E. & JURKOVIC, D. 2007. Value of preoperative ultrasound examination in the selection of women with adnexal masses for laparoscopic surgery. *Ultrasound in Obstetrics and Gynaecology*, 30, 883-888.

ZAGZEBSKI, J. A. 1996. *Essentials of ultrasound physics*, Missouri, Mosby inc.

ZIMMER, Y., TEPPER, R. & AKSELROD, S. 2003. An automatic approach for morphological analysis and malignancy evaluation of ovarian masses using B-scans. *Ultrasound in Medicine & Biology*, 29, 1561-1570.

ZOOROB, R., ANDERSON, R., CEFALU, C. & SIDANI, M. 2001. Cancer screening guidelines. *American Family Physician*, 63, 1101-1112.

APPENDICES

APPENDIX I

Toshiba Aplio 500 specification

Item		Specifications
Power	Line voltage	230 VAC +/-10% (for Europe) 110 to 120 VAC +/-10% (for Other 1) 220 to 240 VAC +/-10% (for Other 2)
	Line frequency	50 Hz to 60 Hz +/-1Hz
	Power consumption	Maximum 1.5 kVA
Operating environmental conditions	Ambient temperature	10 ⁰ C to 30 ⁰ C
	Relative humidity	35% to 80% (no condensation)
	Atmospheric pressure	700 hPa to 1060 hPa
Storage and transportation conditions	Ambient temperature	-10 ⁰ C to 50 ⁰ C
	Relative humidity	30% to 80% (no condensation) 50% or less if ambient temperature exceeds 40 ⁰ C
	Atmospheric pressure	700 hPa to 1060 hPa
External dimensions (not including optional units)	For system with CRT monitor	540 +/-20 mm (width) x 1455 +/-30 to 1500 +/-50 mm (height) x 814 +/-30 mm (depth)
	For system with LCD monitor	540 +/-20 mm (width) x 1360 +/-30 to 1595 +/-50 mm (height) x 814 +/-30 mm (depth)
Mass (not including optional units)	For system with CRT monitor	150 kg or less
	For system with LCD monitor	140 kg or less

APPENDIX II



GIG
CYMRU
NHS
WALES

Bwrdd Iechyd Prifysgol
Caerdydd a'r Fro
Cardiff and Vale
University Health Board

Ysbyty Athrofaol Cymru
University Hospital of Wales

Heath Park,
Cardiff, CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

Parc Y Mynydd Bychan,
Caerdydd, CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 2074 3838
Minicom 029 2074 3632

Eich cyf/Your ref
Ein cyf/Our ref
Webb Health Telephone Network 1872
Direct line/Llinell uniongyrchol

Tel: 029 20746986
Fax: 029 20745311
CAV_Research.Development@wales.nhs.uk

From: Professor JI Bisson
R&D Director
R&D Office, 2nd floor TB2
University Hospital of Wales
Cardiff
CF14 4XW

26 April 2013

Rana Al-Dahlawi
Medical Physics and Clinical Engineering
University Hospital of Wales
Cardiff and Vale UHB
Heath Park
Cardiff

Dear Ms Al-Dahlawi

Cardiff and Vale UHB Ref: 12/OTD5512 : Image Texture Analysis Of Transvaginal Ultrasound In Diagnosing Ovarian Cancer

Thank you for your recent correspondence addressing the reviewers' comments on the above project. Your response was reviewed by the Chair of the Cardiff and Vale Research Review Service (CaRRS).

The Panel is now satisfied with the scientific quality of your proposal, and I can confirm that the following documents have received favourable scientific review:

Document	Version	Date
Protocol	-	-

You may now apply for review by an NHS Research Ethics Committee and NHS R&D governance review.

Please follow the application instructions below:

[For NHS REC review](#)

Please contact your sponsor (Cardiff University RACD) to obtain the sponsor's representative signature needed on your IRAS REC form prior to your submission to the NHS Research Ethics Committee. This signature can be requested electronically using the authorisations tab on your IRAS application.

Version 2.0, 12-4-13

Page 1 of 2

Bwrdd Iechyd Prifysgol Caerdydd a'r Fro yw enw gweithredol Bwrdd Iechyd Lleol Prifysgol Caerdydd a'r Fro
Cardiff and Vale University Health Board is the operational name of Cardiff and Vale University Local Health Board



UHW102X

For NHS R&D approval

Site Specific Information (SSI) form: Question 23 now requires confirmation that you have discussed your project with the R&D Office prior to submission. This allows identification of any issues which might cause a delay in your application and any local authorisations which will be required. Please telephone the R&D Office on 029 20745472 prior to ticking the declaration in question 23.

R&D form: Please obtain the sponsor's representative signature from Cardiff University RACD. This signature can be requested electronically using the authorisations tab on your IRAS application.

Please submit the completed IRAS NHS R&D form and SSI form and all supporting study documentation to NISCHR PCU who will coordinate completion of governance checks prior to R&D permission being granted.

Please ensure that you notify NISCHR PCU if any changes to your protocol or study documents are required by the Research Ethics Committee in order to obtain a favourable ethical opinion.

Final R&D permission to begin your study in Cardiff & Vale UHB will be issued following completion of the governance review by Cardiff and Vale UHB and NISCHR PCU.

YOU SHOULD NOT BEGIN YOUR PROJECT BEFORE RECEIVING WRITTEN CONFIRMATION OF NHS R&D PERMISSION TO BEGIN.

If you require any further information or assistance, please do not hesitate to contact the staff in the R&D Office.

Yours sincerely,



Professor Jonathan I Bisson
Chair of the Cardiff and Vale Research Review Service (CaRRS)

CC: Philippa Young, R&D Lead for Radiology
Sponsor contact, Helen Falconer, Cardiff University
Chief Investigator, Professor Neil Pugh

Link: 'Gaining NHS research permission from Cardiff and Vale UHB – Guidance for researchers' <http://www.cardiffandvaleuhb.wales.nhs.uk/opendoc/180875>

APPENDIX III

Part of the research infrastructure for Wales funded by the National Institute for Social Care and Health Research, Welsh Government.
Yn rhan o seilwaith ymchwil Cymru a ariannir gan y Setyldad Cenedlaethol ac gyfer Ymchwil Gofal Cymdeithasol ac Iechyd, Llywodraeth Cymru



Dyfed Powys Research Ethics Committee

Postal address: PO Box 108
Building 1
St David's Park
Carmarthen SA31 3WY
(for sat nav SA31 3HB)

Telephone : 01267 225045

Fax : 01267 225226

E-mail : sue.byng@wales.nhs.uk

Website : www.nres.nhs.uk

Professor Neil Pugh
Consultant Clinical Scientist
Cardiff & Vale University Health Board
Medical Physics and Clinical Engineering Department
Heath park
UHW
CF14 4XW

18 July 2013

Dear Professor Pugh

Study title: Image texture analysis of ultrasound in diagnosing ovarian cancer
REC reference: 13/WA/0206
Protocol number: SPON 1193-13
IRAS project ID: 124900

Amendments were received on 15 July 2013, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information was considered in correspondence by a sub-committee of the REC. A list of the sub-committee members is attached.

We plan to publish your research summary wording for the above study on the NRES website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to withhold permission to publish, please contact the Co-ordinator Mrs Sue Byng, sue.byng@wales.nhs.uk.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

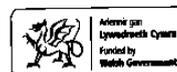
Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).



Cynhelir Cytweithrediad Gwyddor Iechyd Academaidd y Sefylliad Cenedlaethol ac gyfer Ymchwil Gofal Cymdeithasol ac Iechyd gan Fwrdd Addysgu Iechyd Powys
The National Institute for Social Care and Health Research Academic Health Science Collaboration is hosted by Powys Teaching Health Board



APPENDIX IV



Bwrdd Iechyd Prifysgol
Caerdydd a'r Fro
Cardiff and Vale
University Health Board

Ysbyty Athrofaol Cymru
University Hospital of Wales

Parc y Mynydd Bychan
Caerdydd CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Heath Park
Cardiff CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

Image Texture Analysis of Transvaginal Ultrasound

in diagnosing Ovarian Cancer

9 July 2013

Information Sheet (version 1.1)

We would like to invite you to take part in a research study. This is being undertaken as part of a PhD qualification. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish.

(Part 1 tells you the purpose of this study and what will happen to you if you take part).

(Part 2 gives you more detailed information about the conduct of the study).

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

Ultrasound is usually the first examination you will have to assess a mass in your pelvis. However, reading the ultrasound image is very difficult. Cancers sometimes look like non-cancers on the ultrasound image and non-cancers like cancers. Therefore some cancers can be missed.

We want to test a new method of reading the ultrasound images by computer to see if we can improve the accuracy of diagnosing cancer of the ovaries, before you go to surgery. We will compare this method with the techniques already used to see if it is better.

Why have I been invited?

You have been invited to take part in this study because you have been shown to have a cyst on one of your ovaries. If you are pregnant, you should inform us and should not take part in this study.

Do I have to take part?

It is up to you to decide. We will describe the study and go through this information sheet, which we will then give to you. We will then ask you to sign a consent form to show you have agreed to take part. You are free to withdraw at any time, without giving a reason. This would not affect the standard of care you receive.

Bwrdd Iechyd Prifysgol Caerdydd a'r Fro yw enw gweithredol Bwrdd Iechyd Lleol Prifysgol Caerdydd a'r Fro.
Cardiff and Vale University Health Board is the operational name of Cardiff and Vale University Local Health Board





GIG
CYMRU
NHS
WALES

Bwrdd Iechyd Prifysgol
Caerdydd a'r Fro
Cardiff and Vale
University Health Board

Ysbyty Athrofaol Cymru University Hospital of Wales

Parc y Mynydd Bychan
Caerdydd CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Health Park
Cardiff CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

What will happen to me if I take part?

If you agree to take part you will have a second internal ultrasound scan using a more sophisticated ultrasound machine. The scan will be performed in the Doppler Ultrasound Unit at the University Hospital of Wales, Cardiff. Travel expenses to the hospital will not be reimbursed. The scan will be exactly the same as the first scan you had, but this machine automatically collects and stores the data we need. The scan will take about the same time as your first scan (about 15 minutes). The data will be analysed later with our new software. Some of your medical information will be taken from your file. There is nothing else you need to do. If you are booked for surgery through your doctor, the results of our scan will be compared with the results taken from your tissue samples.

What are the possible disadvantages and risks of taking part?

Like any other routine internal scan you may feel minimal discomfort. This scan uses ultrasound and this does not pose any known risks to your health.

What are the possible benefits of taking part?

This study will not affect your treatment, but the information we get from this study may help improve the diagnosis of women with ovarian cancer.

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

This completes part 1.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.





**GIG
CYMRU
NHS
WALES**

Bwrdd Iechyd Prifysgol
Caerdydd a'r Fro
Cardiff and Vale
University Health Board

**Ysbyty Athrofaol Cymru
University Hospital of Wales**

Parc y Mynydd Bychan
Caerdydd CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Heath Park
Cardiff CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

Part 2 of the information sheet

What will happen if I don't want to carry on with the study?

You can withdraw from the study at any time. We cannot use incomplete data, so any data that has been collected will be deleted if you withdraw. The treatment of your pelvic mass will not be affected in any way.

What if there is a problem?

If you have a concern about any aspect of this study you should ask to speak to the researcher who will do her best to answer your questions. If you have any concerns after completing the study, you can contact the researcher by telephoning 02920 743547, or e-mailing al-dahlawirh@cf.ac.uk. If you remain unhappy, and wish to complain formally, you can do this by writing to the Complaints Manager, Cardiff and Vale University Health Board, University Hospital of Wales, Heath Park, Cardiff, CF14 4XW, or by telephoning 02920 746296.

In the event that something does go wrong, and you are harmed during the research, and this is due to someone's negligence, then you may have grounds for a legal action for compensation against the Cardiff and Vale University Health Board, but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Will my taking part in this study be kept confidential?

If you join the study, some parts of your medical records and the data collected for the study will be anonymised and looked at by authorised persons participating in the research. They may also be looked at by representatives of regulatory authorities and by authorised people to check that the study is being carried out correctly. All will have a duty of confidentiality to you as a research participant and we will do our best to meet this duty. No identifiable data will be retained or analysed.

Involvement of the General Practitioner/Family doctor (GP).

This study will in no way affect your health or healthcare, therefore we will not be informing your GP of your involvement. However, if you have any concerns, you are free to contact your GP yourself, and your GP may contact us for information, or to answer any questions.

Bwrdd Iechyd Prifysgol Caerdydd a'r Fro yw enw gweithredol Bwrdd Iechyd Lled Prifysgol Caerdydd a'r Fro. Cardiff and Vale University Health Board is the operational name of Cardiff and Vale University Local Health Board.





GIG
CYMRU
NHS
WALES

Bwrdd Iechyd Prifysgol
Caerdydd a'r Fro
Cardiff and Vale
University Health Board

Ysbyty Athrofaol Cymru
University Hospital of Wales

Parc y Mynydd Bychan
Caerdydd CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Heath Park
Cardiff CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

What will happen to the results of the research study?

The result of the research will be published in peer viewed medical journals, but we assure you none of the participant will be identified in any report/publication.

Who is organising and funding the research?

This study has been organised and funded by Cardiff University in conjunction with UHW.

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by Dyfed Powis Research Ethics Committee.

Who can I contact for further information about the research and contact details?

If you have any concerns during the study you can contact the research head,

Mrs Rana Aldahlawi,

Department of Medical Physics and Clinical Engineering,

Telephone: 029 2074 3547

For further information about research you can look up online websites:

National Electronic Library for Health <http://www.library.nhs.uk/trials>

The National Research Register - UK database of research projects <http://www.nrr.nhs.uk/>

CancerHelp UK - There is a search to help people find cancer clinical trials and trial information.

<http://www.cancerhelp.org.uk>

<http://www.cancerhelp.org.uk/help/default.asp?page=51>

The National Translational Cancer Research Network - A list of trials for people with cancer that are currently recruiting patients. www.ntrac.org.uk

Bwrdd Iechyd Prifysgol Caerdydd a'r Fro yw enw gweithredol Bwrdd Iechyd Usol Prifysgol Caerdydd a'r Fro
Cardiff and Vale University Health Board is the operational name of Cardiff and Vale University Local Health Board



APPENDIX V



Bwrdd Iechyd Prifysgol
Caerdydd a'r Fro
Cardiff and Vale
University Health Board

Ysbyty Athrofaol Cymru
University Hospital of Wales

Parc y Mynydd Bychan
Caerdydd CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Heath Park
Cardiff CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

Consent Form version 1.1

Image Texture analysis of Transvaginal Ultrasound in diagnosing

Ovarian Cancer

9 July 2013

Patient identification number _____

Researcher: Rana AL-Dahlawi

Please
Initial
box

1. I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
2. I understand that my participation is voluntary, and that I am free to withdraw at any time without giving any reason, without my medical care or legal right being affected.
3. I confirm that I am not pregnant or taking part in any other medical research study.
4. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from Cardiff University, from regulatory authorities or from the NHS trust, where it is relevant to my taking part in this research. I give permission to these individuals to have access to my records.
5. I understand that I will undergo an additional internal ultrasound scan as part of the research. I understand that this will not be used in any diagnoses or to inform any medical decision made about me.
6. I agree to take part in the above study.

Name of patient

Date

Signature

Name of person taking consent

Date

Signature

Bwrdd Iechyd Prifysgol Caerdydd a'r Fro yw enw gweithredol Bwrdd Iechyd Lleol Prifysgol Caerdydd a'r Fro.
Cardiff and Vale University Health Board is the operational name of Cardiff and Vale University Local Health Board



- BOOTE, E. J. 2003. Doppler US Techniques: Concepts of Blood Flow Detection and Flow Dynamics. *Radiographics*, 23, 1315-1327.
- BROSKY, J. 2009. Toshiba ASQ delivers hard data on liver fibrosis. Available: [www.european-hospital.com/en/article/6676-Toshiba ASQ delivers hard data on liver fibrosis](http://www.european-hospital.com/en/article/6676-Toshiba_ASQ_delivers_hard_data_on_liver_fibrosis) [Accessed 27-02-2014].
- DEPRIEST, P. D., SHENSON, D., FRIED, A., HUNTER, J. E., ANDREWS, S. J., GALLION, H. H., PAVLIK, E. J., KRYSZCIO, R. J. & VAN NAGELL JR, J. R. 1993. A morphology index based on sonographic findings in ovarian cancer. *Gynecologic Oncology*, 51, 7-11.
- DONG, L., CUI, H., LI, X. P., SUN, L. F., CHANG, X. H., LIANG, X. D. & ZHU, H. L. 2008. Clinical value of serum CA19-9, CA125 and CP2 in mucinous ovarian tumor: a retrospective study of 273 patients. *Zhonghua fu chan ke za zhi*, 43, 5-8.
- HAMID, B. 2011. *The Reliability of B-mode transvaginal probe image for the quantitative texture analysis and the dependence of extracted features on region of interest size for ovarian cancer detection*. PhD, Cardiff University.
- HORSKIN, P. 2010. *Principle of Doppler Ultrasound*, Cambridge: Cambridge University Press.
- KAIJSER, J., SAYASNEH, A., VAN HOORDE, K., GHAEM-MAGHAMI, S., BOURNE, T., TIMMERMAN, D. & VAN CALSTER, B. 2014. Presurgical diagnosis of adnexal tumours using mathematical models and scoring systems: A systematic review and meta-analysis. *Human Reproduction Update*, 20, 449-462.
- LU, R. S., TIAN, G. Y., GLEDHILL, D. & WARD, S. 2006. Grinding surface roughness measurement based on the co-occurrence matrix of speckle pattern texture. *Applied Optics*, 45, 8839-8847.
- LUAN, N.-N., WU, Q.-J., GONG, T.-T., VOGTMANN, E., WANG, Y.-L. & LIN, B. 2013. Breastfeeding and ovarian cancer risk: a meta-analysis of epidemiologic studies. *American journal of clinical Nutrition*, 113.
- MEDICAL PHYSICS, W. 2009 (online). *Acoustic structure quantification ramps echo resolution* [Online]. Available: <http://medicalphysicsweb.org/cws/article/newsfeed/40669> [Accessed 27-02-2014].
- MODAN, B., HARTGE, P., HIRSH-YECHEZKEL, G., CHETRIT, A., LUBIN, F. & WACHOLDER, S. 2001. Parity, Oral Contraceptives, and the Risk of Ovarian Cancer among Carriers and Noncarriers of a BRCA1 or BRCA2 Mutation. *The New England Journal of Medicine* 235-240.
- PRYS DAVIES, A., JACOBS, I., WOOLAS, R., FISH, A. & ORAM, D. 1993. The adnexal mass: Benign or malignant? Evaluation of a risk of malignancy index. *British Journal of Obstetrics and Gynaecology*, 100, 927-931.
- ROSSING, M., DALING, J. & WEISS, N. 2013. Ovarian tumors in a cohort of infertile women. *The New England Journal of Medicine*, 331.
- RUBIN, J. M. 1994. Spectral Doppler US. *Radiographics : a review publication of the Radiological Society of North America, Inc*, 14, 139-150.
- RUMACK, C. (ed.) 2005. *Diagnostic Ultrasound*: Mosby inc.
- RUMACK, C. 2011. *Diagnostic Ultrasound*, Mosby inc.
- SZCZYPIŃSKI, P. M., STRZELECKI, M., MATERKA, A. & KLEPACZKO, A. 2009. MaZda—A software package for image texture analysis. *Computer Methods and Programs in Biomedicine*, 94, 66-76.

