The Value of Ultrasound for the Diagnosis and Management of Ovarian Tumours

By
Joseph Yazbek

The Early Pregnancy and Gynaecological Assessment Unit
King's College Hospital, London

Thesis submitted for the Degree of Doctor of Medicine

Postgraduate Medical School
University of Surrey

August 2009

© Joseph Yazbek 2009
Abstract

Objectives:
The aim of this thesis was to evaluate the value of modern high-resolution transvaginal ultrasound for differential diagnosis of adnexal tumours. This was done by performing a series of studies: (1) comparison of risk of malignancy index (RMI) and ovarian crescent sign (OCS) for the diagnosis of ovarian malignancy; (2) assessment of the value of pattern recognition for the diagnosis of borderline ovarian tumours (BOT); (3) comparison of diagnostic accuracy using real-time and static images (4) assessment of the reproducibility of ultrasound pattern recognition for the diagnosis of BOT; (5) assessment of the effect of degree of operator’s confidence on diagnostic accuracy; (6) examination of the effect of the quality of gynaecological ultrasound diagnosis on the management of patients with suspected ovarian cancer; and (7) determining the value of preoperative ultrasound examination for selection of women for laparoscopic surgery.

Interpretation of findings:
(1) The OCS proved to be a simpler and better method in diagnosing ovarian cancer when compared to the RMI. (2) Ultrasound diagnosis of BOT is highly specific but typical features are absent in one third of cases. (3) The diagnosis of an adnexal mass is more accurate when made on the basis of real-time ultrasound examination than on static ultrasound images. (4) Accuracy of ultrasound diagnosis of BOT is less in comparison to benign and invasive malignant lesions. (5) Accuracy of ultrasound pattern recognition in differentiating ovarian tumours depends on the degree of certainty with which the diagnosis is made. (6) Improved quality of ultrasound examination has a measurable positive effect on the management of women with suspected ovarian malignancy. (7) Detailed
preoperative ultrasound examination of benign adnexal lesions is helpful for assessing feasibility of their removal by laparoscopic surgery.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE</td>
<td>1</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>2</td>
</tr>
<tr>
<td>TABLE OF CONTENT</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>11</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>15</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>17</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>18</td>
</tr>
<tr>
<td>DECLARATION</td>
<td>19</td>
</tr>
<tr>
<td>PART ONE: BACKGROUND</td>
<td>20</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>21-28</td>
</tr>
<tr>
<td>CHAPTER 2: ULTRASOUND GRAY-SCALE MORPHOLOGY (PATTERN RECOGNITION)</td>
<td>29-37</td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td></td>
</tr>
<tr>
<td>2.2 Lesion</td>
<td></td>
</tr>
<tr>
<td>2.3 Locularity</td>
<td></td>
</tr>
<tr>
<td>2.4 Papillary projections</td>
<td></td>
</tr>
<tr>
<td>2.5 Cyst wall</td>
<td></td>
</tr>
<tr>
<td>2.6 Cyst content</td>
<td></td>
</tr>
<tr>
<td>2.7 Solid lesion</td>
<td></td>
</tr>
<tr>
<td>2.8 Acoustic shadows</td>
<td></td>
</tr>
<tr>
<td>2.9 Ascites</td>
<td></td>
</tr>
</tbody>
</table>
2.10 Size

2.11 Qualitative classification of adnexal lesions

2.12 'Simple' cysts

2.13 Complex adnexal cysts

CHAPTER 3: OTHER METHODS USED IN THE ULTRASOUND ASSESSMENT OF ADNEXAL PATHOLOGY

3.1 Introduction

3.2 Colour Doppler

3.3 Morphological Scoring Systems

3.4 Multiparameter Models
   3.4.1 Risk of Malignancy Index
   3.4.2 Logistic Regression Models
   3.4.3 Artificial Neural Network Models

CHAPTER 4: SERUM CA-125

CHAPTER 5: OVARIAN TUMOURS

5.1 Introduction

5.2 Demographics of Ovarian Cancer

5.3 Natural History of Ovarian Cancer

5.4 Benign Ovarian Tumours
   5.4.1 Follicular cysts
   5.4.2 Corpus Luteum cysts
   5.4.3 Theca-Lutein cysts
   5.4.4 Serous inclusion cysts
   5.4.5 Endometriomas
5.4.6 Epithelial tumours
5.4.7 Germ cell tumours
5.4.8 Sex cord-stromal tumours
5.4.9 Other Benign Adnexal tumours
  5.4.9.1 Hydro-pyo-haemato-salpinx
  5.4.9.2 Tubo-ovarian abscess
  5.4.9.3 Paraovarian cysts
  5.4.9.4 Peritoneal pseudocysts
5.5 Borderline Ovarian Tumours
5.6 Malignant ovarian tumours
  5.6.1 Epithelial ovarian cancers
  5.6.2 Sex-cord stromal ovarian cancers
  5.6.3 Germ cell ovarian cancers
  5.6.4 Metastatic ovarian tumours
5.7 Management of Ovarian Cancer

PART TWO: AIMS, METHODS AND RESULTS 79
CHAPTER 6: AIMS 80
6.1 Aims of the thesis

CHAPTER 7: MATERIALS, METHODS AND RESULTS 81-88
7.1 Setting
7.2 Gynaecology population
7.3 Preoperative assessments
7.4 Randomisation
7.5 Ultrasonography
7.6 Colour Doppler Imaging
7.7 Serum CA 125 measurements
7.8 Histopathology and staging
7.9 Statistical Analysis
7.10 Ethics Committee approval

CHAPTER 8: STUDY 1  
A COMPARATIVE STUDY OF THE RISK OF MALIGNANCY INDEX AND THE OVARIAN CRESCENT SIGN FOR THE DIAGNOSIS OF INVASIVE OVARIAN CANCER

8.1 Background
8.2 Methods
8.3 Statistical Analysis
8.4 Results

CHAPTER 9: STUDY 2  
THE ACCURACY OF ULTRASOUND SUBJECTIVE “PATTERN RECOGNITION” FOR THE DIAGNOSIS OF BORDERLINE OVARIAN TUMOURS

9.1 Background
9.2 Methods
9.3 Statistical Analysis
9.4 Results
CHAPTER 10: STUDY 3  
REAL-TIME ULTRASOUND VERSUS EVALUATION OF STATIC IMAGES IN THE  
PREOPERATIVE EVALUATION OF ADNEXAL MASSES  
10.1 Background  
10.2 Methods  
10.3 Statistical Analysis  
10.4 Results  

CHAPTER 11: STUDY 4  
THE USE OF ULTRASOUND PATTERN RECOGNITION BY EXPERT  
ULTRASOUND OPERATORS TO IDENTIFY BORDERLINE OVARIAN TUMOURS:  
A STUDY OF DIAGNOSTIC PERFORMANCE AND INTEROBSERVER  
AGREEMENT OF ULTRASOUND DIAGNOSES  
11.1 Background  
11.2 Methods  
11.3 Statistical Analysis  
11.4 Results  

CHAPTER 12: STUDY 5  
CONFIDENCE OF THE ULTRASOUND OPERATOR IN MAKING A DIFFERENTIAL  
DIAGNOSIS OF AN ADNEXAL TUMOUR: EFFECT ON DIAGNOSTIC ACCURACY  
AND INTEROBSERVER AGREEMENT  
12.1 Background  
12.2 Methods  
12.3 Statistical Analysis  
12.4 Results
CHAPTER 13: STUDY 6  141-157

EFFECT OF THE QUALITY OF GYNAECOLOGICAL ULTRASONOGRAPHY ON MANAGEMENT OF PATIENTS WITH SUSPECTED OVARIAN CANCER: A RANDOMISED CONTROLLED TRIAL

13.1 Background
13.2 Methods
13.3 Statistical Analysis
13.4 Results

CHAPTER 14: STUDY 7  158-167

VALUE OF PREOPERATIVE ULTRASOUND IN THE SELECTION OF WOMEN WITH ADNEXAL MASSES FOR LAPAROSCOPIC SURGERY

14.1 Background
14.2 Methods
14.3 Statistical Analysis
14.4 Results

PART THREE: DISCUSSIONS  168

CHAPTER 15: STUDY 1  169-171
CHAPTER 16: STUDY 2  172-176
CHAPTER 17: STUDY 3  177-178
CHAPTER 18: STUDY 4  179-181
CHAPTER 19: STUDY 5  182-183
CHAPTER 20: STUDY 6  184-187
CHAPTER 21: STUDY 7  188-192
LIST OF TABLES

CHAPTER 5

1. International Federation of Gynecology and Obstetrics (FIGO) staging of primary ovarian carcinoma. 78

CHAPTER 8

2. Types of ovarian lesions and management strategies applied in each group. 95
3. Histology and International Federation of Gynaecology and Obstetrics stage of invasive ovarian cancer. 96
4. False-positive cases in the diagnosis of invasive ovarian cancer by the risk of malignancy index (RMI) and ovarian crescent sign (OCS). 97

CHAPTER 9

5. Demographic data and clinical symptoms at presentation of 166 women with an adnexal mass. 106
6. Histology and International Federation of Gynaecology and Obstetrics stages of borderline and invasive ovarian cancers. 107
7. Gray-scale ultrasound characteristics and histology of the tumours of the study population. 108
8. Ultrasound morphological appearance of false positive and false negative cases. 109
9. Age of women and tumour volume in true positive, false positive and false negative cases of borderline ovarian tumour (BOT). 110
10. Accuracy of pattern recognition for the diagnosis of different types of ovarian tumours.

CHAPTER 10

11. Histopathological diagnoses of the adnexal masses included in the study.

12. Accuracy, sensitivity, and specificity with regard to malignancy of subjective evaluation of gray-scale and Doppler ultrasound findings in an adnexal mass during scanning ("real-time" sonologist), and by three "image experts" (A, B, C) who evaluated the static images saved by the real-time sonologist, and of the "consensus opinion" of the three image experts (i.e., the diagnosis suggested by at least two of the three experts).

CHAPTER 11

13. The diagnosis of individual experts in benign, borderline and primary invasive ovarian tumours.

14. Diagnostic performance of pattern recognition with regard to discriminating between invasive and non-invasive (benign and borderline) tumours.

15. Diagnostic performance of pattern recognition with regard to discrimination between malignant (borderline and invasive malignancies) and benign tumours.

16. Interobserver agreement between each two of the three experts in classifying the tumours as benign, borderline and invasive malignant (N = 166).
17. Proportion of experts who correctly classified the ovarian masses. 130

CHAPTER 12
18. Diagnostic confidence of the three operators. 135
19. Diagnostic accuracy according to operator's confidence (N = 166). 136
20. Agreement in diagnosis versus agreement in confidence between the operators (N=166). 137
21. Interobserver agreement in classifying the tumours as benign, borderline and invasive malignant between each two of the three experts, stratified by agreement in confidence in making the diagnosis (certain and probable). 138
22. Rate of agreement in both diagnosis and confidence between the operators, stratified by histology (N = 166). 139
23. Cases misdiagnosed by all three operators. 140

CHAPTER 13
24. Demographic data, source of referral and presenting symptoms. 153
25. Management of all patients and final histopathological findings from 71 patients. 154
26. Histology and International Federation of Gynaecology and Obstetrics stage of malignant adnexal tumours. 155
27. Outcome of the patients managed expectantly. 156
28. Accuracy of ultrasonography for the diagnosis of malignant adnexal tumours. Data included only patients in whom preoperative diagnosis from ultrasonography was conclusive of nature of the adnexal tumour and for whom histological diagnosis was also available. 157
CHAPTER 14

29. Demographic features of women and adnexal tumour characteristics of women booked for laparotomy or laparoscopy.

30. Histological subtypes of adnexal tumours operated on by laparoscopy and laparotomy.
LIST OF FIGURES

CHAPTER 5
1. Number of new cases and age specific incidence rates, female ovarian cancer, UK 2004 59

CHAPTER 8
2. A complex ovarian cyst with healthy ovarian tissue (arrows) clearly visible adjacent to the wall of the cyst (the ovarian crescent sign). 98
3. Receiver–operating characteristics curve comparing the diagnostic performance of the risk of malignancy index (RMI) and the ovarian crescent sign (OCS) for the diagnosis of invasive ovarian cancer. Solid line, RMI >200; Dashed line, OCS. 99

CHAPTER 9
4. Ultrasound image showing a serous borderline ovarian tumour with extensive papillary projections arising from the cyst wall. Note the presence of healthy ovarian tissue, the ‘ovarian crescent sign’ (arrow). 112
5. Ultrasound image showing a mucinous endocervical-type borderline ovarian tumour (BOT). Note the resemblance to a serous BOT, but more organized papillary projections. The ovarian crescent sign is also a common feature (arrow). 113
6. Ultrasound image showing a mucinous gastrointestinal-type borderline ovarian tumour. A ‘honeycomb nodule’ is seen suspended within the cyst cavity (arrow). 114
CHAPTER 13
7. Flow diagram of study participants. US, ultrasound; CT, computed tomography.

CHAPTER 14
8. Flow chart of the patients in the study.
<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial neural networks</td>
</tr>
<tr>
<td>BOT</td>
<td>Borderline ovarian tumour</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>IOTA</td>
<td>International ovarian tumour analysis</td>
</tr>
<tr>
<td>LR-</td>
<td>Negative likelihood ratio</td>
</tr>
<tr>
<td>LR+</td>
<td>Positive likelihood ratio</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>NPV</td>
<td>Negative predictive value</td>
</tr>
<tr>
<td>NS</td>
<td>Not significant</td>
</tr>
<tr>
<td>OCS</td>
<td>Ovarian crescent sign</td>
</tr>
<tr>
<td>PMB</td>
<td>Postmenopausal bleeding</td>
</tr>
<tr>
<td>PPC</td>
<td>Primary peritoneal cancer</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>PI</td>
<td>Pulsatility index</td>
</tr>
<tr>
<td>RI</td>
<td>Resistance index</td>
</tr>
<tr>
<td>RMI</td>
<td>Risk of malignancy index</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver-operating characteristic</td>
</tr>
<tr>
<td>RR</td>
<td>Relative risk</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>TAMXV</td>
<td>Time averaged maximum velocity</td>
</tr>
<tr>
<td>TAS</td>
<td>Transabdominal ultrasound scanning</td>
</tr>
<tr>
<td>TVS</td>
<td>Transvaginal ultrasound scanning</td>
</tr>
<tr>
<td>US</td>
<td>Ultrasound</td>
</tr>
</tbody>
</table>
Acknowledgments

I would like to extend my heartfelt gratitude to my teacher and supervisor, Mr. Davor Jurkovic (Consultant Gynaecologist, University College Hospital, London) for his continued guidance, generosity and support during my years as a Research Fellow and beyond. His help, stimulating suggestions and encouragement helped me in all the time of research and for the writing of this thesis. I am extremely grateful to him for his teaching in scanning and for helping me to develop professionally in the field of Obstetrics and Gynaecology.

I am indebted to Miss Shanti Raju (Consultant Gynaecological Oncologist, South East London Gynaecological Cancer Centre, Guy’s and St Thomas’ Foundation Trust, London) for her vital encouragement and support and for her invaluable advice in the majority of studies.

My former colleagues from the Department of Obstetrics and Gynaecology at King’s College Hospital and at Guy’s and St Thomas’ Hospital, London. I want to thank them for all their help, support, interest and valuable hints. Especially I am obliged to Dr Jara Ben Nagi, Dr Emma Sawyer, Dr Chris Lee and Dr Samir Helmy for their help in recruiting patients for the studies included in this MD thesis.

Finally, I am grateful to all the members of the gynaecological oncology team at Guy’s and St Thomas’ NHS Foundation Trust and the staff at the Early Pregnancy and Gynaecological Assessment Unit at King’s College Hospital Foundation Trust for their help with the conduct of the studies included in this thesis. I am also grateful to all the ultrasonographers who contributed towards the randomised study.
Declaration

The work conducted in this thesis was carried out in the Early Pregnancy and Gynaecological Assessment unit of King’s College Hospital, London and in Guy’s and St Thomas’ NHS Foundation Trust, London. I was personally involved in the design of all the studies and recruitment of patients. I coordinated day-to-day liaison between the two study sites, checked, coded, and entered data, and obtained additional information where data were missing. The studies had local ethical committee approval and all patients included gave informed consent.
PART ONE:

Background
Chapter 1: Introduction

The reported prevalence of ovarian tumours has significantly increased in the past two decades (Borgfeldt and Andolf, 1999; Nelson et al., 1986; Yazbek et al., 2007). This increase is mainly attributed to the routine use of ultrasound for the assessment of women with a wide range of gynaecological complaints. In many cases women with clinical suspicion of adnexal mass are urgently referred by their general practitioners to a rapid access unit and they usually receive a scan as a part of the investigations. In many women the diagnosis of ovarian cyst generates considerable anxiety because of uncertainty about its nature. The presence of an ovarian cyst is traditionally considered to be an indication for operative intervention due to fear of ovarian cancer and acute complications of ovarian cysts such as torsion and cyst rupture. However, these risks are difficult to quantify and it is possible that many asymptomatic women with cysts are being over-treated. Surgery itself is not without risks and many asymptomatic women with benign cyst may develop complications following an operation to remove the ovarian cyst (Chapron et al., 1998; Minelli, 1996).

The management of patients diagnosed with ovarian tumours is mainly determined by their nature. The best survival rates for women with ovarian cancer are achieved when treatment is organised and carried out by gynaecological oncologists working in cancer centres (Junor et al., 1999). However, women with asymptomatic benign tumours can be managed expectantly (Yazbek et al., 2007; Platek et al., 1995, Lee et al., 2004) or by minimally invasive surgery if symptomatic (Hilger et al., 2006; Mais et al., 1995). These operations can be safely performed by general gynaecologists in their local hospitals.
There is an added pressure on the ultrasonographers to provide a diagnostic opinion about the nature of the adnexal mass being benign or malignant. Furthermore, there is still no consensus on the most effective diagnostic criteria to discriminate between benign and malignant ovarian tumours on ultrasound scan (Granberg et al., 1990; Sassone et al., 1991; Lerner et al., 1994; Jacobs et al., 1990; Tailor et al., 1997; Timmerman et al., 1999a; Valentin et al., 2001; Aslam et al., 2000). Ultrasound methods used to establish a diagnosis in the case of an adnexal lesion will be discussed in further detail in the relevant chapters.

Most ultrasonographers in the United Kingdom use the pattern recognition method in order to establish a diagnosis of an ovarian tumour. However, the differentiation between benign and malignant adnexal tumours is often difficult and it mainly depends on the expertise of the ultrasonographers (Valentin, 1999; Timmerman et al., 1999b). Timmerman et al. (1999b) aimed at differentiating benign and malignant ovarian tumours using subjective pattern recognition in a study of 300 patients, taking into account the influence of observer’s experience on overall test performance. The best reported accuracy (calculated as the sum of true positives and true negatives divided by the total number of studied cases) was achieved by the most experienced operator and this was 275/300 (91.7 %, 95% CI 88 – 94.3). The study was performed on stored still ultrasound images and did not classify borderline ovarian tumours separately from invasive malignant tumours.

Valentin (1999) assessed the value of ultrasound pattern recognition method in 173 cases scheduled for surgery because of pelvic masses. She reported a sensitivity of 88% (95% CI, 69 – 96), specificity 96% (95% CI, 92 – 98), positive likelihood ratio 22 and negative likelihood ratio 0.13. Again this study classified borderline ovarian tumours as invasive malignant.
Tumours with complex morphological appearance on ultrasound such as borderline ovarian tumours (BOT) are difficult to characterise and they are easily confused with invasive ovarian malignancy (Valentin et al., 2006). The inability to establish accurately the nature of the ovarian tumour would eventually lead the gynaecologist to perform additional tests such as checking serum CA-125 as this could be used to establish a risk of malignancy index (RMI) (Jacobs et al., 1990) based on which results, the patient could be offered management locally or could be referred to a gynaecological cancer centre for management (RCOG Greentop guidelines, 2003).

Jacobs et al. (1990) assessed the risk of malignancy index on 143 patients preoperatively. A cut-off level of 200 provided a sensitivity of 85.4% (95% CI, 70.8 – 94.4), specificity 96.9 (95% CI, 91.3 – 99.4), positive likelihood ratio 42.1 and negative likelihood ratio 0.15. The problem is that the study included a small proportion of ovarian malignancies (n=42), which were mainly stage III and of serous histology. This could have contributed to the high CA-125 level, which is only elevated in 50% of stage I ovarian cancer.

Some centres, have facilities to refer those patients with inconclusive scan result to a tertiary gynaecological ultrasound centre, where the scans are carried out by ‘experts’ in gynaecological ultrasound scan (EFSUMB Newsletter, 2005). Experienced ultrasonographers using relevant clinical information and their subjective assessment of ultrasonographic images can differentiate malignant from benign masses in most cases (Timmerman et al., 1999b). However, it is unclear whether the gynaecologist and the patient would be completely reassured and opt for conservative management if a diagnosis of benign adnexal tumour is made by experts in gynaecological ultrasonography. Furthermore, the reproducibility of ultrasound pattern recognition, when used by different
experts in gynaecological ultrasonography, for the differentiation between benign, borderline and malignant invasive ovarian tumours is not known. We also do not know whether expert ultrasound operators use similar criteria to differentiate between the three different types of adnexal tumours and how confident these operators are in reaching their final diagnoses. In addition, an important issue in gynaecological ultrasonography yet to be investigated and that is whether expert ultrasound operators could review still ultrasound images, when their opinion is required and they are not personally available to scan the patient, and achieve a high diagnostic accuracy (in our thesis diagnostic accuracy is expressed through sensitivity and specificity, positive and negative predictive values, or positive and negative diagnostic likelihood ratios) comparable to when they are performing the real-time ultrasonography.

Despite the fact that the diagnostic accuracy of expert ultrasound operators in differentiating between benign and malignant tumours is high (Valentin, 1999; Timmerman et al., 1999b), these experts are not available in all hospitals and the day to day ultrasound examinations are carried out by ultrasonographers or by general gynaecologists. Thus, it was important to describe “easy” ultrasound markers, which may help less experienced ultrasonographers in the differentiation between benign and malignant adnexal tumours. The RMI, which has been widely adopted in the UK to facilitate triage of women with ovarian tumours for referral to tertiary gynaecological oncology units (Jacobs et al., 1990; RCOG Greentop guidelines, 2003). Although the RMI is relatively a simple test to use in clinical practice, it has significant false negative and false positive rates (Aslam et al., 2000). In addition, the results are not immediately available and the test is relatively costly, as it involves the use of serum biochemistry. Our research group have recently described a new morphological ultrasound feature, the ‘ovarian crescent sign’ (OCS), which depends
on the fact that healthy ovarian tissue can be seen adjacent to the cyst within the ipsilateral ovary (Hillaby et al., 2004). The initial report, which included 100 women, reported a sensitivity of 91% (95% CI, 76 – 97), specificity 84% (95% CI, 73 – 91), positive likelihood ratio 5.54 and negative likelihood ratio 0.11 in the diagnosis of ovarian malignancy (this included borderline ovarian tumours). This study was not applied on the general population but only included women who were scheduled to have surgery to remove their ovarian tumours. A relatively low proportion of benign ovarian tumours (n=67) was included as a result. This study showed that this morphological ultrasound sign has a potential of becoming a simple and effective way of excluding an invasive ovarian malignancy, without the need for a detailed morphological assessment of the tumour or the use of serum biochemistry. We carried out further work on the OCS and prospectively compared its value to that of RMI for the diagnosis of ovarian cancer in a large group of women with an ultrasound diagnosis of ovarian tumour.

BOTs remain the most difficult adnexal tumours to be correctly classified both pre-operatively and on histological examination (Valentin et al., 2006). Since 1971, the International Federation of Gynecology and Obstetrics (FIGO) (1971) has classified BOTs as an entity separate from benign and invasive ovarian tumours (Serov et al., 1973). However, until recently, BOTs have been treated in the same way as invasive malignant tumours. BOTs are more prevalent in women of childbearing age (Gotlieb et al., 2005, Gotlieb et al., 1998; Boran et al., 2005) and their prognosis is generally good (Ahmed and Lawton, 2005; Silverberg et al., 2004). They often follow a relatively benign course (Ahmed and Lawton, 2005; Silverberg et al., 2004) and therefore fertility sparing conservative surgery is often contemplated in women with BOT who wish to preserve their reproductive potential (Maneo et al., 2004). Therefore, accurate diagnosis is essential for
planning appropriate patient management (Osmers, 1996). The diagnostic accuracy of ultrasound pattern recognition for the differential diagnosis between benign, borderline and invasive ovarian tumours has not been tested as yet. We collected data on a selected mix of women who were referred to our gynaecological ultrasound assessment unit with adnexal tumours, which nature was “difficult” to characterise. We prospectively assessed the value of pattern recognition for the differential diagnosis of adnexal tumours and in particular its accuracy in establishing the specific diagnosis of histological subtypes of BOTs.

In our hospitals, a patient with an adnexal mass is often first scanned by a junior or senior sonographer. Subsequently, the images are discussed with a supervising ultrasound operator, who will often rescan the patient, because he or she believes that this will help in gaining the right information to make a correct diagnosis. However, to the best of our knowledge, it has never been confirmed in a scientific study that real-time scanning is superior to evaluating saved static images of an adnexal mass. If examining static ultrasound images had a comparable accuracy to real-time ultrasonography, this could have a positive effect on the management of patients in small hospitals where expertise could be lacking and a second opinion might be sought by sending the images to an expert ultrasound operator in a tertiary hospital. We examined a large data set containing representative images of different types of adnexal tumours in order to determine if the prediction of the character of an adnexal mass using subjective evaluation of gray-scale and Doppler ultrasound findings based on the evaluation of static images is as accurate as that based on a real-time ultrasound examination.

Previous studies have shown that the ultrasound pattern recognition method is superior to other ultrasonographic methods for differentiating between benign and malignant adnexal tumours (Valentin, 1999), especially if the examination was carried out
by an experienced ultrasound operator (Timmerman et al., 1999b). However, the reproducibility of expert ultrasound operators at identifying borderline ovarian tumours as a separate entity from benign and malignant invasive ovarian tumours has not been properly tested before. We carried out a study to assess reproducibility of ultrasound pattern recognition method for the diagnosis of borderline ovarian tumours by asking experienced ultrasound operators to systematically examine a large data set containing representative images of different types of adnexal tumours.

Ultrasound pattern recognition technique is operator-dependent and its accuracy improves with the increase in the operator’s experience. However, it is unclear whether expert ultrasound operators have a similar degree of confidence in making their diagnoses and whether the diagnostic accuracy decreases with the decrease in the level of confidence. We used the same dataset from the above study to assess the degree of confidence with which expert ultrasound operators differentiate benign, borderline and malignant invasive ovarian tumours and to measure the diagnostic accuracy and interobserver agreement when the ultrasound operators make a certain, probable or uncertain diagnoses.

In recent years we have witnessed significant improvements in the ability of ultrasound examination to discriminate between different types of adnexal tumours. It is not clear, however, whether this improved diagnostic accuracy has a measurable impact on the management of women with ovarian tumours. We carried out a prospective randomised study to check whether the improved specificity of ultrasound diagnosis would lead to a significant decrease in the number of major surgical procedures in women who were referred to our regional gynaecological cancer centre with suspected ovarian cancer. This study also investigated whether, and for the first time, the efforts to improve the quality of ultrasound diagnosis are worthwhile. The study assessed the role of highly skilled
ultrasound operators in the management protocols of women diagnosed with adnexal tumours as this may have important benefits: shorter and more effective diagnostic assessment, reduced number of major surgical procedures and their associated morbidity, better use of subspecialty oncology services and reduced cost of treatment.

In the past, several studies have assessed the value of ultrasound and Doppler in correctly identifying the risk of malignancy of adnexal tumours and selecting those with low risk of malignancy for laparoscopic surgery (Berlanda et al., 2002; Guerriero et al., 2005). However, the value of ultrasound for the preoperative selection of women with benign adnexal masses for removal by either laparoscopic surgery or by laparotomy has not been examined. We collected prospective data on women who had non-invasive adnexal tumours diagnosed on ultrasound scan and required surgery. These women underwent a pre-operative scan by a highly skilled ultrasound operator in order to develop and assess a set of ultrasound criteria that were designed to select women for laparoscopic removal of ovarian cyst with a minimum risk of converting the operation to a laparotomy. Although there is no doubt that the success of laparoscopic surgery is mainly dependent on the skill and expertise of the operating surgeon. In expert hands most operations can be completed successfully using a laparoscopic approach. Most consultant gynaecologists in the UK and worldwide perform intermediate-level laparoscopic surgery (RCOG in collaboration with BSGE, 2005). The results of this study may aid them when selecting women with benign cysts for minimally invasive surgery.

In summary, this work investigated whether the improved diagnostic accuracy of gynaecological ultrasonography has a measurable impact on the management of women with ovarian tumours. It also aimed at aiding less experienced ultrasound operators in correctly differentiating between different types of adnexal tumours.
2.1. Introduction

Ovarian morphology differs and depends on the woman's age and the menopausal status. It is very difficult to measure the ovaries of a female child (Orsini et al., 1984; Sample et al., 1977). On the other hand, the premenopausal ovarian morphology is a continually varying landscape, which is controlled by the hypothalamic-pituitary axis. Thus the ovarian morphology would differ depending on the day of the menstrual cycle. Whereas, the inability to visualise one or both ovaries in postmenopausal women is not unusual due to the relative inactivity of these ovaries.

The introduction of transvaginal ultrasound scanning (TVS) into practice provided ultrasound operators with better resolution in order to assess the morphology of the normal and the abnormal ovary. TVS has a high patient acceptability (Van Nagell Jr et al., 1991). Transabdominal ultrasound scanning (TAS) has limited resolution and requires the patient to have a full bladder in order to provide an acoustic window. Nonetheless, TAS and TVS complement each other in order to ensure complete examination of large ovarian tumours, extending beyond the range of the transvaginal transducer.

There was a lack of standardised description of findings in gynaecological sonography (Timmerman et al., 2000). A review of the literature had revealed considerable variation in the diagnostic accuracy of test procedures as a result of this lack of standardisation (Timmerman, 1997). Consequently, a new initiative was started by the International Ovarian Tumor Analysis (IOTA) group, where a steering committee had made recommendations about terms and procedures for morphologic end-points by B-mode imaging (Timmerman et al., 2000).
2.2 Lesion

An adnexal lesion is the part of an ovary or an adnexal mass that is judged from the assessment of ultrasound images to be inconsistent with normal physiologic function. This situation can arise from the presence of a persistent unilocular cyst, surrounded by healthy ovarian tissue. In this case the whole ovary containing the cyst is the ‘ovary’, whilst the unilocular cyst is the ‘lesion’. The size of both structures is measured independently, and the cyst is described as being unilocular (see below). In cases where the lesion is separate from the ovaries (e.g. hydrosalpinx), the size of both ovaries and the lesion are measured separately. When no normal ovarian stroma is visible; the lesion and the ovary are indistinguishable, and consequently the size of the lesion and the ovary will be the same (Timmerman et al., 2000).

2.3 Locularity

A loculated cyst is one that is divided into compartment by septa. A septum is defined as a thin strand of tissue running across the cyst cavity from one internal surface to the contralateral side. An incomplete septum is defined as a thin strand of tissue running across the cyst cavity from one internal surface to the contralateral side, but which is not complete in some scanning planes. When the septum is incomplete the cyst is described as unilocular.

The number of compartment in cystic ovarian tumours is related to the potential likelihood of the lesion being malignant. A previous study by Osmers et al. (1996) found that in the absence of solid components, the clinical significance of a tumour increases in the presence of more than three compartments within the cyst. A study by Meire et al. (1978) found that multilocular cysts are more likely to be malignant when compared to unilocular cysts.
The septum thickness has been analysed in previous studies. Malignant tumours are more likely to have thick and irregular septa compared to thin septa in benign lesions. A thick septum has been arbitrarily defined as one greater than 3 mm. Meire et al. (1978) found the majority of ovarian cysts with thick septa were malignant.

2.4. Papillary projections

These are solid projections arising from the cyst wall into the cyst cavity with a height greater than or equal to 3 mm. These solid tissues are thought to represent localised overgrowths of epithelium. Although they are seen occasionally in benign ovarian tumours, they appear to be characteristic features of both borderline ovarian tumours and stage I epithelial ovarian cancer (Valentin et al., 2006a). It is also believed that the greater the number of papillary projections present, the greater the likelihood of malignancy (Granberg et al., 1990; Bailey et al., 1998).

It is occasionally difficult to report solid papillary projections in the presence of incomplete septum. It is recommended (Timmerman et al., 2000) that excrescences associated with the 'cogwheel’ sign and the 'beads-on-a-string’ sign (usually observed in hydrosalpinges) (Timor-Tritsch et al., 1998) should be classified as papillary projections if their height is greater or equal to 3 mm. It is also important not to classify the hyper-reflective area in the centre of a dermoid cyst (Cohen and Sabbagha, 1993) as a solid papillary projection (Timmerman et al., 2000). In this thesis a projection with a thin base (<2 mm in width) especially an isolated one was recorded as an incomplete septum.
2.5 Cyst wall

The internal cyst wall is described as being smooth or irregular. In the presence of at least one solid papillary projection, then the wall is described as irregular. The external wall of the cyst is not taken into account. The risk of malignancy is greater with the increase in the internal wall irregularity.

2.6 Cyst content

The dominant feature of the tumour content is described as anechoic when the content appears black like in the case of functional cysts. The content may alternatively be described as low-level echogenic as seen in mucinous tumours, 'ground glass' appearance when the homogenously dispersed echogenic cystic contents is noted like in the case of endometriotic cysts, haemorrhagic appearance (star-shaped, cobweb-like, or jelly-like), or mixed echogenicity like in the case of teratomas. Solid tumours tend to have high echogenicity, however the dominant feature of any cystic contents is only described if it can be assessed.

2.7 Solid lesion

Solid tumours exhibit high echogenicity suggesting the presence of tissue like in the case of ovarian fibromas. Occasionally, cystic tumours containing thick fluid could be mistaken for solid tumours like in the case a blood clot filling an ovarian cyst. Pushing the transducer gently towards the structure and looking for internal movement may distinguish between blood clots and solid tissue. However in cases of uncertainty, it is recommended to classify the lesion as solid.
2.8 Acoustic shadows

This is defined as the loss of acoustic echo behind a sound-absorbing structure. Acoustic shadowing is noted in the presence of small calcified particles within the tumour like in the case of dermoid cysts or ovarian fibromas.

2.9 Ascites

This is the presence of free fluid outside the pouch of Douglas, within the peritoneal cavity. It is usually documented as present or absent. Ascites is typically present in cases of ovarian malignancies. Osmers et al. (1996) reported the presence of free fluid in 11.1% of borderline ovarian tumours, but free fluid was also present in 13% of cases with functional cysts. In a more recent study, Valentin et al. (2006b) reported the presence of ascites in 9% of borderline ovarian tumours. This was less, statistically, than the frequency of ascites in the case of Stage I epithelial ovarian cancer (31%) and Stage II-IV epithelial ovarian cancer (61%). This group also observed the presence of ascites in cases of rare ovarian malignancies (56%) and in cases of metastatic ovarian tumours (43%).

2.10 Size

An adnexal tumour is measured in three perpendicular planes and its mean diameter is calculated according to the formula \( \frac{(D_1 + D_2 + D_3)}{3} \). The tumour volume is calculated using the formula \( \frac{\pi}{6} \times D_1 \times D_2 \times D_3 \).

Several studies have demonstrated that the risk of malignancy increases with the size of the lesion (Granberg et al., 1990; Andolf et al., 1986). Granberg et al. (1990) reported a positive predictive value for malignancy of 5.9% for cysts less than 5 cm in mean diameter, 21.3% from 5 to 10 cm, and 43.6% for cysts greater than 10 cm. Osmers et
al. (1998) reported an incidence of malignancy of 13.9% in cysts less than 4 cm in mean diameter compared to an incidence of malignancy of more than 50% in cysts greater than 7 cm in mean diameter.

In another study Osmers et al. (1996) found that 36.9% of all premenopausal ovarian tumours to be between 3 and 4 cm in diameter. Of these 68.2% turned out to be functional in nature. They also found that 27.4% of tumours greater than 9 cm in mean diameter were malignant.

Rulin and Preston (1987) found, in a study population of 150 postmenopausal women with adnexal pathology, that only one of 47 malignant masses to be less than 5 cm in diameter. In addition, only one of 32 masses less than 5 cm in diameter was found to be malignant. The likelihood of malignancy in postmenopausal women with adnexal masses less than 5 cm in diameter was 3% (Rulin and Preston, 1987).

2.11 Qualitative classification of adnexal lesions

Timmerman et al. (2000) suggested classifying all lesions into one of six categories:

- Unilocular cyst: A cyst without septa or solid parts or papillary projections.
- Unilocular solid cyst: a unilocular cyst with a measurable solid component or at least one papillary projection. Hydro- or pyosalpinges may be included in this category, if the height of the ‘beads-on-a-string’ is greater than or equal to 3 mm.
- Multilocular cyst: a cyst with at least one septum but with no measurable solid components or papillary projections.
- Multilocular solid cyst: a multilocular cyst with a measurable solid component or at least one papillary projection.
• Solid tumour: a tumour where the solid components comprise 80% or more of the tumour, when assessed in a two dimensional section
• Not classifiable: owing to poor visualisation

2.12 ‘Simple’ cysts

These are thin, smooth walled unilocular cysts with anechoic fluid content, with no visible septum or solid components. The minimum diameter of simple cysts used in all the studies forming this thesis was 3 cm.

The risk of malignancy within simple cysts less than 10 cm in diameter is minimal (Auslender et al., 1996; Levine, 1994; Bailey C et al., 1988). However, when a cyst diameter approaches and extends beyond the range of the transvaginal probe (usually up to 8–10 cm) a complete survey of the internal wall is not possible, and “simple cyst” classification should be used with caution in these large ovarian cysts. In these cases, transvaginal ultrasound and transabdominal ultrasound should be subsequently used in order to ensure complete examination of the lesion. These simple cysts may represent physiological (functional) cysts, serous cystadenomas, or peritoneal inclusion cyst.

Osmers et al. (1996) estimated that around 60% of all premenopausal ovarian tumours are simple cysts. Two-thirds of these are functional cysts that would resolve spontaneously. This group found that 0.8% of ovarian malignancies presented as simple cyst (Osmers et al., 1996). The risk of malignancy would obviously increase with the diameter of the simple cyst.

The prevalence of simple cysts in postmenopausal women is reported to be between 3% and 17% (Auslender et al., 1996; Levine, 1994). The most common histology of simple cysts in this group of women is benign cystadenoma. The risk of malignancy is small when
the cyst is less than 5 cm in diameter (Roman, 1998). Ekerhovd et al. (2001) compared preoperative sonographic morphology with gross and histological findings in 927 pre- and 337 postmenopausal women with unilocular cysts. They found three borderline and three malignant tumours at surgery in supposedly simple cysts, all of which were over 7.5 cm in greatest diameter. All had visible papillations at surgery. In addition, there was one paradoxically benign truly simple ovarian cyst of 3.9 cm associated with stage III cancer and ascites. The authors calculated the false-positive rate as being seven of 660 simple cysts or 1.06%, including the benign cyst, in all women. There were no cancers in this series of simple cysts, which were less than or equal to two centimetres. Of the 377 total unilocular (both simple and complex) postmenopausal cysts, almost a third were benign tumours, 28% were histologically simple cysts, 6% were functional cysts, 6% were hydrosalpinges, 8% were endometriomas and 7% were paraovarian cysts. There were five tumours of low malignant potential and 12 cancers.

Valentin and Akrawi (2002) methodically observed 134 asymptomatic women with 160 benign appearing adnexal cysts found incidentally and followed over 8 years. The findings of this study reinforced the concept that simple cysts do not require surgery.

However, some borderline ovarian tumours (BOT) may exhibit a unilocular smooth-walled appearance. 9% of the BOTs described by Exacoustos et al. (2005) and 4% of those described by Fruscella et al. (2005) were unilocular smooth-walled cysts. This again could be due to the fact that this type of ovarian tumour is characterized by its large size at presentation, which makes the complete assessment of the entire tumour a difficult task.
2.13 Complex adnexal cysts

These are multilocular cysts, which may or may not contain echogenic fluid or solid papillary projections. It has been shown that the more complex the tumour appearance, the greater is the likelihood of it being malignant (Granberg et al., 1990). However, the term complex cyst could represent any adnexal tumour ranging from a functional corpus luteum cyst to an invasive ovarian malignancy. Therefore a large degree of overlap between benign and malignant pathology exists in this group of adnexal tumours. Thus, it is recommended to use the standardised terms in describing the appearances of different adnexal tumours as detailed in section 2.11.

Subjective evaluation of the gray-scale ultrasound image, i.e. pattern recognition, for discrimination between benign and malignant tumours can almost certainly be learnt by anyone performing gynaecological ultrasound examinations on a regular basis, but diagnostic accuracy increases with increasing experience (Timmerman et al., 1999b). An experienced ultrasound examiner can very confidently discriminate between benign and malignant pelvic tumours in the adnexal region using pattern recognition, the reported sensitivity of pattern recognition varying between 96 and 98%, specificity between 89 and 90%, LR+ between 8.72 and 9.8 and LR- between 0.02 and 0.05 (Timmerman et al., 1999b).

Pattern recognition has been shown to be superior to all other ultrasound methods (these will discussed separately in further chapters) (Valentin, 1999; Jain, 1994; Buy et al., 1996; Levine et al., 1994; Salem et al., 1994; Stein et al., 1999; Aslam et al., 2000a; Valentin et al., 2001).
3.1 Introduction

Assessment of an adnexal mass can be made using B-mode imaging alone. Colour Doppler studies have been also introduced in order to help in the differentiation between benign and malignant adnexal tumours. It was hoped that the introduction of Doppler ultrasound would allow detection of abnormal angiogenesis within the pelvis and this would subsequently lead to better detection of malignant changes within a mass.

Morphological scoring systems and statistical models, such as those involving multivariate logistic regression analysis and more recently artificial neural networks, have been developed, to facilitate the differentiation of benign from malignant, and in the case of mathematical models, to assign an individual “risk” of malignancy. They aim to try and mimic the processes of evaluation that occur within the mind of an experienced ultrasonographer, in assigning a risk of malignancy when presented with a number of parameters concerning a given patient or lesion (Tailor et al, 1997; Timmerman et al., 1999a). Considerable overlap exists between the individual ultrasonographic findings that are used to define the malignant and benign features of a mass. Initial studies set out to establish that the presence or absence of certain morphological characteristics was predictive of ovarian malignancy. Increasing complexity, in the form of locularity of the cyst, or irregularities of the cyst wall, corresponded with an increasing probability of the lesion being malignant.
3.2 Colour Doppler

Colour Doppler ultrasound of adnexal lesions helps to identify vascularised tissue and can assist in differentiating solid tumour tissue from non-vascularised structures. It is also used in conjunction with pulsed Doppler US to identify vessels for waveform analysis. Areas of blood flow can be seen on Doppler ultrasound as areas of colour.

The Doppler information obtained is colour frequency-coded i.e. colour is assigned based on the direction of blood flow and the variance of detected frequency shifts. Red indicates blood flow towards the probe, and the blue indicates blood flow away from the transducer. The brightness of the colour is proportional to the velocity of flow within the vessel.

The gain and pulse repetition frequency are initially adjusted for maximum sensitivity of low blood flow. The lowest velocity signals are filtered out gradually by increasing the pulse repetition frequency and flow analysis is concentrated on the highest velocity signals. A pulsed Doppler range gate is placed over the area of interest to provide flow velocity waveforms that can be analysed. The indices measured are the pulsatility index (PI), resistance index (RI), peak systolic velocity (PSV) and time-averaged maximum velocity (TAMXV). RI and PI value represent a measure of impedance to blood flow distal to the point of sampling (Thompson et al., 1988). A low value indicates decreased impedance to blood flow, whereas a high value is indicative of increased impedance to blood flow.

Most studies have relied on waveform analysis to distinguish benign from malignant ovarian masses (Kurjak and Predanic, 1992; Franchi et al., 1995; Kurjak et al., 1992; Brown et al., 1998). Benign lesions tend to initiate new tumour blood vessel formation peripherally from pre-existing host vessels, whereas malignant tumours tend to
initiate new tumour blood vessel formation centrally (Kurjak et al., 1993; Dock et al., 1991; Folkman et al., 1989). Waveform analysis is based on the fact that malignant tumour vessels are morphologically abnormal: they lack smooth muscle in their walls and demonstrate an irregular course and arteriovenous shunt formation (Kurjak et al., 1993; Dock et al., 1991; Folkman et al., 1989). In addition, malignant tumour vessels generally have low impedance, which causes high diastolic flow and low systolic-diastolic variation. Some differentiation between benign and malignant masses is achieved by quantifying these differences.

A comparison of different studies shows that no standard has been established concerning which Doppler index to use or what cut-off value is most appropriate. However, RI less than 0.4–0.8 (84–90) and PI less than 1.0 are generally considered to be suspicious for malignancy (Hamper et al., 1993; Brown et al., 1998; Jain, 1994; Levine et al., 1994; Salem et al., 1994). Doppler ultrasound has yielded variable results in distinguishing benign from malignant masses, with a sensitivity of 50%–100% and a specificity of 46%–100% (Hata et al., 1992; Kurjak et al., 1992; Reles et al., 1997; Timor-Tritsch et al., 1993). Differing results are partly due to varying threshold values and corresponding tradeoffs between sensitivity and specificity.

In cases in which septations, papillae, and solid areas of tumour are absent, it is difficult to detect signal for waveform analysis. In addition, certain Doppler indexes can be misleading in premenopausal women and usually have a lower specificity because of physiologic alterations in the ovary due to the menstrual cycle which cause lower blood vessel resistance, thereby mimicking malignancy. Finally, acute inflammatory adnexal disease and endometriosis are common conditions associated with an increased number of capillaries and dilatation of blood vessels, which cause a low PI (Pellerito et al., 1995). In a
study by Reles et al. (1997), the sensitivity of colour Doppler ultrasound was 80%, and the specificity only 67% (LR+ of 2.42 and LR- of 0.30) in premenopausal patients, whereas in postmenopausal patients, the sensitivity and specificity were 93% and 83%, respectively with LR+ of 5.47 and LR- of 0.08.

Valentin et al. (1994) compared the ability of colour Doppler and that of gray-scale ultrasound morphology to distinguish between benign and malignant adnexal tumours. They found that PI, peak systolic volume (PSV) and TAMXV values overlapped in almost all benign and malignant masses, except for multilocular cysts with solid parts. In this group of cysts the PI was significantly lower, and the blood flow velocities significantly higher, in malignant tumours. The ultrasound morphology was a better discriminator between benign and malignant tumours compared to Doppler variables. They concluded that Doppler has a limited role in the assessment of adnexal pathology.

In summary, variations exist in the reported results of Doppler ultrasound. At present, there are no reliable criteria regarding absolute cut-off values for the different Doppler indices in the discrimination of benign and malignant adnexal tumours. Adding Doppler ultrasound examination to subjective evaluation of the gray-scale ultrasound image does not seem to yield much improvement in diagnostic precision but it may increase the confidence with which a correct diagnosis of benignity or malignancy is made (Valentin, 2004).
3.3 Morphological Scoring Systems

Because the diagnostic performance of ultrasound pattern recognition method is operator dependent (Valentin 1999; Timmerman et al., 1999b) and in order to overcome these limitations, scoring systems were designed with an aim to define strict criteria when describing ultrasound findings, which were easily reproducible.

Morphological scoring systems were introduced in order to improve the diagnostic accuracy of ultrasound in the differentiation between benign and malignant adnexal masses. Granberg et al. in 1989 demonstrated a correlation between the ultrasonographic appearance and the microscopic characteristics of a cyst. Within the population studied, 296 cysts were characterised as simple and unilocular. Of these, only one was malignant, and this cyst was between 5–10 cm in diameter, with a macroscopically visible papillary process in a woman aged 60 years. Of the cysts with papillations 93% were malignant, and papillations were the most predictive morphological feature of malignancy, when compared to the woman's age or cyst size. There was no correlation between the size of a simple cyst and malignancy, however, increasing cyst size in complex cysts will correlate with an increased risk of malignancy. It must also be remembered that a cyst size of greater than 10 cm in diameter may be difficult to characterize as a full assessment of the internal cyst wall will not always be possible due to the limited scanning field of the transvaginal probe. An extension of initial morphological classifications was the development of scoring systems to assign a risk of malignancy to a particular mass.

Sassone et al. (1991) developed one of the most popular scoring systems to allow an objective description of extrauterine pelvic disease, based on morphological characteristics. They were able to differentiate benign from malignant tumours with a specificity of 83%, sensitivity of 100%, LR+ of 5.88 and LR- of 0. The positive and negative predictive powers
were 37 and 100%, respectively. The variables studied were inner wall structure, wall thickness, septal thickness, and the echogenicity of cyst contents. Tumour size did not improve the sensitivity, and so it was excluded. Benign cystic teratomas were the main cause of false positive test results.

Lerner et al. (1994) modified the morphological scoring system devised by Sassone et al. (1991) by producing a weighted scoring system including the presence or absence of acoustic shadowing, resulting in better differentiation between benign and malignant masses. The sensitivity and specificity achieved were 96.8% and 77% respectively (LR+ of 4.21 and LR- of 0.04), with a positive predictive value of 29.4%. They concluded that further amendments to the scoring system were necessary in order to reduce the high false positive rates. They postulated whether the addition of Doppler flow studies would improve the diagnostic performance (Lerner et al., 1994).

Bourne et al. in 1993 prospectively tested the use of a morphology score as a second stage test within an ovarian cancer screening program. Its use, while significantly reducing the false positive rate, also reduced the sensitivity to 83% from 100%.

3.4 Multiparameter Models

These models were introduced in an attempt to improve sensitivity and specificity in the ultrasound diagnosis of adnexal tumours. These models combine demographic, biochemical and ultrasound data to calculate a probability of malignancy.
3.4.1 Risk of Malignancy Index

Described by Jacobs et al. in 1990, this was one of the first models, which combined ultrasound, demographic and biochemical data. This risk of malignancy index has been widely adopted in the UK to facilitate triage of women with ovarian tumours for referral to tertiary gynaecological oncology units (Jacobs et al., 1990; RCOG Greentop guidelines, 2003).

Jacobs et al. (1990) included one hundred and forty-three women who were scheduled to have surgery for an adnexal mass in their study. Demographic data were recorded for each patient and all women had a serum CA-125 level checked pre-operatively. 139 women underwent a pre-operative transabdominal scan. Postmenopausal status was defined as greater than one year of amenorrhoea or age over 50 in the case of prior hysterectomy.

The study included 101 benign cases and 42 malignant cases. Stepwise logistic regression analysis of the data revealed that menopausal status, ultrasound score and serum CA-125 levels were the most significant variables for determining the risk of malignancy.

Using logistic regression analysis Jacobs et al. (1990) devised the following risk of malignancy index (RMI):

\[ RMI = U \times M \times \text{serum CA-125 (kU/L)} \]

Each of the following gray-scale morphological features was given one point when present: bilateral lesions, multilocular lesions, solid areas, intra-abdominal metastases and ascites. If the sum of points was 1, an ultrasound score \( U = 1 \) was given, while a sum of \( \geq 2 \) was given a score \( U = 3 \). Premenopausal women were given a score \( M = 1 \) and postmenopausal women were given a score \( M = 3 \).
By analysing their results, Jacobs et al. (1990) found that using a RMI cut-off level of 200 will achieve the greatest diagnostic accuracy with a sensitivity of 85.4% (95% CI, 70.8 – 94.4), specificity 96.9 (95% CI, 91.3 – 99.4), LR+ of 42.1 and LR- of 0.15.

Davies et al. (1993) tested the RMI proposed by Jacobs et al. (1990) using the same cut-off value of 200, they reached a sensitivity of 87%, specificity of 89%, LR+ of 7.9 and LR- of 0.15 in diagnosing ovarian malignancy.

Tingulstad et al. (1996) evaluated the RMI devised by Jacobs et al. (1990) and proposed a second RMI (RMI 2). The RMI 2 was calculated using the following formula:

\[ \text{RMI 2} = U \times M \times \text{serum CA125 (kU/L)}. \]

If the sum of points was 0 or 1, an ultrasound score \( U = 1 \) was given, while a sum of \( \geq 2 \) was given a score \( U = 4 \). Premenopausal women were given a score \( M=1 \) and postmenopausal women were given a score \( M=4 \).

RMI 2 had better sensitivity compared to RMI 1 (80% vs. 71%) but slightly worse specificity (92% vs. 96%).

The problem is that CA-125 level is a main component of the RMI. However, this tumour marker may be elevated in several benign conditions, which mainly affect premenopausal women such as menstruation, endometriosis or recent ovulation. As a result, the RMI is likely to have lower specificity in premenopausal women despite the fact that the menopausal status score is lower (M=1) in this group of women.
3.4.2 Logistic Regression Models

These are mathematical models based on multivariate logistic regression analysis, with the aim of assigning an individual risk of malignancy. These are based on a statistically derived weighting system of certain ultrasonographic and demographic characteristics, rather than the arbitrary definitions of more simple morphology scores. This allows the relationship of each individual factor to be related to each other, as well as the histological diagnosis. In order to use mathematical models within routine clinical practice, a large initial data set is required.

Tailor et al. (1997) used a multivariate logistic regression model to assign a probability of malignancy to sixty-seven women undergoing surgery for adnexal masses. All women underwent a transvaginal B-mode and colour Doppler imaging. Ten variables were recorded for each patient. Only three of these variables (age of patient, papillary projection score and the time averaged maximum velocity (TAMXV)) were considered significant and were retained in the final equation. The papillary projection score was assigned values of either 1 or 0 indicating their presence or absence respectively in transvaginal ultrasound assessment. The dataset included 52 women with benign, three with borderline and 12 with invasive ovarian tumours.

The probability of malignancy was calculated using the following formula:

\[
\frac{1}{1 + e^{-z}}
\]

Where \(z = (0.1273 \times \text{Age}) + (0.2794 \times \text{TAMXV}) + (4.4136 \times \text{papillary projection score}) - 14.2046.\)

At a cut-off value of 25%, they achieved the best sensitivity and specificity of 93.3% and 90.4% respectively. They concluded that prediction of malignancy using their
Timmerman et al. (2005) described a logistic regression model based on the assessment of 1066 women with adnexal masses (800 benign and 266 malignant). All patients underwent transvaginal ultrasonography with colour Doppler imaging.

The most useful variables for the logistic regression analysis were: personal history of ovarian cancer, hormonal therapy, age, maximum diameter of the lesion, pain, ascites, blood flow within a solid papillary projection, presence of an entirely solid tumour, maximal diameter of solid component, irregular internal cyst walls, acoustic shadows and a colour score of intratumoural blood flow. The model containing all the 12 variables gave an area under the receiver operating characteristic curve of 0.95 for the development data set (n=754 patients). The corresponding value for the test data set (n=312 patients) was 0.94, and a probability cut-off value of .10 gave a sensitivity of 93%, specificity of 76%, LR+ of 3.88, LR- of 0.09, PPV of 55% and NPV of 97%.

3.4.3 Artificial Neural Network Models

Artificial neural networks (ANN) are able to investigate non-linear relationships between variables, as opposed to logistic regression models, which are based on linear classification. ANN are essentially computer-based decision making tools, which are modelled on the structure and learning behaviour of human biological nervous systems (Shepherd and Koch, 1990; Haykin, 1994; Newey, 1997).

Knowledge is acquired by the network through a learning process (i.e. training) and, second, interneuron connection strengths (known as synaptic weights) are used to store the knowledge. These neurons exchange information in the form of numerical values with each
other via synaptic interconnections. The neurons take a weight sum of their inputs, subtract a certain inhibition bias and then feed this numerical value through the network to produce an output. These outputs are then fed into the next neuron layer. Finally, at the output level of the ANN the outcome represents a probability of an adnexal mass being malignant.

A neuron has the same mathematical form as a logistic regression model and therefore an artificial neural network is a generalisation of the logistic regression technique. A unique feature of an ANN is its ability to generalise. Generalisation refers to the neural network producing reasonable outputs for inputs not encountered during training. Neurons within an ANN are structured depending upon the learning algorithm used to train the network. This learning algorithm refers to the procedure used to perform the learning process.

ANN fall into two distinct categories: supervised and unsupervised. Unsupervised networks are capable of partitioning an input feature set (e.g. morphological variables) into distinct subsets with no prior knowledge of subset classes (Newey, 1997). On the other hand, supervised networks require an external teacher to decide for the network, which sets of features belong to the same classes. The adjustment of the network weights (for example, through the back-propagation algorithm) and thus the network's behaviour is based on an external supply of examples of correct performance. In most network models the initial weights are set at random. During training, examples are presented to the network and the synaptic weights have been fine-tuned such that a maximum number of training data were correctly classified. These weight values are then fixed and the network can be used to classify new cases (Gurney, 1997).

Networks may be structured in a variety of different ways (Haykin 1994). Amongst the most commonly used structures for classification are multilayer feedforward networks.
These consist of three layers: an input layer, a hidden layer and output layer. The hidden layer allows the network to extract higher-order statistics (non-linear regression analysis). Single-layer feedforward networks consist of an input layer and an output layer. There is no hidden layer. In its simplest form this is equivalent to linear logistic regression analysis.

The design of a neural network and the selection of the optimal number of source nodes in the input layer and the optimal number of hidden neurons are difficult and results from trial and error. A network that is too complex for the size of the training set is at risk of becoming overtrained and therefore will have a poor generalisation capacity. However, if the network is not complex enough, it will not have a good learning capacity and the test results will be equally suboptimal (Gurney, 1997).

Timmerman et al. (1999a) investigated ANN models from simple clinical and ultrasound-derived criteria to predict whether or not an adnexal mass will have histological evidence of malignancy. They collected prospective data from 173 consecutive patients who were scheduled to undergo surgery for adnexal tumours. The prevalence of malignancy was 28.3% and the best ANN gave a sensitivity of 95.9%, specificity of 93.5%, LR+ of 14.75 and LR- of 0.04.

Tailor et al. (1999) tested ANN models on 67 women with adnexal masses, of which 22.4% were malignant. With an optimum cut-off value of 0.45, they reached a sensitivity of 100%, specificity of 98.1%, LR+ of 52.6 and LR- of 0.

Aslam and Co-workers (2000a) prospectively tested the accuracy of the RMI 1 (Jacobs et al., 1990), RMI 2 (Tingulstad et al., 1996) and Tailor's regression model (Tailor et al., 1997) in the diagnosis of adnexal malignancy. They included 61 patients in their study. Of these, 38 women had benign tumours and 23 had ovarian cancer. Tailor's regression model (Tailor et al., 1997) achieved a sensitivity of 43% (95% CI, 23 – 64),
specificity of 92% (95% CI, 84 – 100), LR+ of 5.38 and LR- of 0.62. This compared with a sensitivity of 74% (95% CI, 56 – 92) and 74% (95% CI, 56 - 92), specificity of 92% (95% CI, 84 – 100) and 89% (95% CI, 80 – 99), LR+ of 9.25 and 6.72 and LR- of 0.28 and 0.29 for RMI 1 and RMI 2 respectively. This study showed that all three models when applied prospectively, they performed less accurately than originally described. The authors (Aslam et al., 2000a) went to speculate that it is likely that the accuracy of these tests would be even less in a population of women in whom there was a substantial clinical uncertainty.
Chapter 4: Serum CA-125

CA-125 antigen is a high molecular weight glycoprotein, which is expressed by a large proportion of epithelial ovarian cancers. It is detected by the OC125 monoclonal antibody, which was first described by Bast et al. in 1981. Serum CA-125 is the most extensively studied ovarian cancer-associated marker. The incorporation of serum CA-125 in the risk of malignancy index (RMI) (Jacobs et al., 1990) and the recommended use of the RMI nationwide as a triaging test for adnexal malignancy (RCOG Greentop guidelines, 2003) has given this tumour marker a very important role and made it widely investigated.

The sensitivity and specificity of CA-125 is known to be poor. It is only raised in approximately 50% of stage 1 epithelial ovarian cancers and in 75–90% of patients with advanced disease. False positive results have been noted in many medical disorders, both malignant (Ozguroglu et al., 1999; Gaspar et al., 2003) and benign (Levine et al., 1997; Meden and Fattahi-Meibodi, 1998; Krishnan et al., 2002; Buamah, 2000).

Examples of benign conditions with raised CA-125 include: endometriosis, ovarian hyperstimulation syndrome, Meig’s and pseudo-Meig’s syndrome, multivisceral tuberculosis, tuberculous peritonitis, liver cirrhosis, fulminant hepatic failure, uraemia and renal failure, nephrotic syndrome and pancreatitis. Malignant conditions with raised CA-125 include ovarian cancer, advanced fallopian tube cancer, advanced uterine cancer, breast cancer with peritoneal metastases, gastric cancer with peritoneal metastases, lymphoma with peritoneal involvement, pancreatic carcinoma, advanced rectal/bladder cancer, and advanced hepatocellular carcinoma (Ahmed and Lawton, 2005). These false positive results are more frequently encountered in premenopausal women. Serum CA-125 level is elevated in 1% of healthy female subjects (Maggino et al., 1994).
Requests for CA-125 testing are increasing, but there are suggestions that a large proportion of this increase is a result of its use by specialties other than gynaecology or oncology, thus leading to concerns that CA-125 is being used by clinicians who are not fully aware of its limitations and its role in ovarian cancer. This may result in missed diagnoses or unnecessary investigation of patients (Moss et al., 2005).

The normal cut-off value for CA-125 is 35 units/ml. These levels are altered in pregnant women (Jacobs et al., 1988) and therefore non-pregnant reference ranges cannot be used for the investigation of cysts detected in pregnancy.

Aslam et al. (2000b) carried out a study on 291 pregnant women, who attended their hospital for the routine first trimester nuchal translucency scan. They attempted to establish a reference range for maternal serum CA-125 levels at 11-14 weeks of gestation. They excluded women who had an ovarian cyst greater than 2 cm in mean diameter and those in whom either ovary was not visualised. They measured the serum CA-125 of 188 women who had a median CA-125 level of 23.4 U/ml (range 2.2-166.3, 95% CI 5.3-70.2). Their data demonstrated that in the investigation of ovarian cysts in pregnancy, the non-pregnancy cut-off values cannot be used. They concluded that if the discriminatory power of serum CA-125 in pregnancy is similar to that in nonpregnant women, a more appropriate cut-off value in pregnancy may be 112 U/ml, which corresponds to the 99th centile in their study (Aslam et al., 2000b).

A multimodality approach (i.e., clinical examination, imaging, and serum assays) is necessary to detect malignant masses at an early stage. The combination of normal findings at serum CA-125 assay, endovaginal US, and clinical examination of the pelvis virtually excludes the possibility of ovarian cancer (Schutter et al., 1994).
Although elevated serum CA-125 levels (>35U/mL) have been found at radioimmunoassay in more than 80% of ovarian cancer patients, only 50% of patients with stage I disease have elevated serum CA-125 levels (Zanaboni et al., 1987; Zurawski et al., 1988; Maggino et al., 1994). Serial measurements of the serum CA-125 level are routine in the clinical follow-up of ovarian cancer patients (Zanaboni et al., 1987).

Increasing serum CA-125 levels in patients with early-stage disease is predictive of recurrence regardless of imaging findings. In 87% of patients, doubling or halving of the serum CA-125 level correlates with tumour progression or regression, respectively (Bast et al., 1983). A combination of serum CA-125 assays and imaging studies may aid in patient treatment. Negative serum CA-125 assays and negative imaging findings do not exclude recurrent tumour; occult tumour will be detected in up to 50% of affected patients at second-look laparotomy (Folk et al., 1995).

Timmerman et al. (2007) investigated the value of serum CA-125 measurements alone or as part of a multimodal strategy to distinguish between malignant and benign ovarian tumours before surgery based on a large prospective multicentre study (International Ovarian Tumor Analysis). 809 Patients with at least one persistent ovarian mass preoperatively underwent transvaginal ultrasonography using gray scale imaging to assess tumour morphology and colour Doppler imaging to obtain indices of blood flow. Of these, 567 (70%) patients had benign tumours and 242 (30%) had malignant tumours. A logistic regression model including CA-125 (M2) resulted in an area under the receiver operating characteristic curve (AUC) of 0.934 and did not outperform a published (M1) without serum CA-125 information (AUC, 0.936). Specifically designed new models including CA-125 for premenopausal women (M3) and for postmenopausal women (M4) did not perform significantly better than the model without CA-125. In postmenopausal
patients, serum CA-125 alone (AUC, 0.920) and the risk of malignancy index (AUC, 0.924) performed very well. Results were very similar when the models were prospectively tested on a group of 345 new patients with adnexal masses of whom 126 had malignant tumours (37%). The authors concluded that adding information on CA-125 to clinical information and ultrasound information does not improve discrimination of mathematical models between benign and malignant adnexal masses (Timmerman et al., 2007).

In summary, CA-125 has a high false positive rate and poor sensitivity and specificity. The substantial inappropriate usage of CA-125 may lead to results that are useless to the clinician, have cost implications, and add to patient anxiety and clinical uncertainty. CA-125 should be used in conjunction with imaging modalities. The patient’s age, menstrual history and other diseases should be taken into consideration when interpreting the CA-125 results.
Chapter 5: Ovarian Tumours

5.1 Introduction

The prevalence of ovarian tumours is on the increase and this could be simply due to the liberal use of transvaginal ultrasound. Clinicians are faced with adnexal tumours that have been diagnosed incidentally. These tumours would not have been detected otherwise before the ultrasound era. The natural history of incidentally detected pelvic tumours with benign ultrasound morphology is not known. Therefore, the optimal management of such tumours is also unknown. It is usually left up to the treating clinician to advise on the best management, although no clear guidelines exist.

Certain adnexal tumours such as functional cysts and hydrosalpinges are best treated expectantly (Valentin, 2004), especially if the patient is asymptomatic. Other tumours like peritoneal pseudocysts may be treated by cyst aspiration (Jain, 2000) and some adnexal tumours require definitive surgery like in the case of borderline or invasive ovarian tumours. It is of prime importance to establish a preoperative diagnosis regarding the nature of an adnexal tumour, because if the tumour is benign then the best treatment would be by removal via laparoscopic surgery (Yuen et al., 1997), whereas malignant masses require to be operated on by laparotomy by a gynaecological oncologist.

This chapter will discuss different types of benign, borderline and malignant adnexal pathology. Ultrasound appearances and management of each tumour will also be discussed. I will also go through the demographics of ovarian cancer and try to shed some light on the natural history of ovarian cancer.
5.2 Demographics of Ovarian Cancer

Ovarian cancer is the most common gynaecological cancer in UK women (ISD Scotland online, 2007; Office for national statistics, 2007; Welsh cancer intelligence and surveillance unit, 2007; Northern Ireland cancer registry, 2007). In 2004 there were 6,615 cases of ovarian cancer diagnosed in the UK. The lifetime risk of developing ovarian cancer is approximately 1 in 48 (2.1%) for women in England and Wales. This risk increases to 2.5% in women with one affected first degree relative and between 30-40% in women with two affected first-degree relatives. Ovarian cancer is predominantly a disease of older, post-menopausal women with almost 85% of cases being diagnosed in women over 50 years of age (5568 out of 6615 cases). There is a steep increase in incidence after the usual age of the menopause. The highest incidence rates are for women aged 65 and over as shown in figure 1. The incidence of ovarian cancer in British women has increased over the last 25-30 years from around 15 per 100,000 women in 1975 to around 17 per 100,000 women in 2004, an increase of more than 16% whilst the mortality rate has remained fairly stable (ISD Scotland online 2007; Welsh cancer intelligence and surveillance unit, 2007; Northern Ireland cancer registry, 2007).

Worldwide there are more than 190,000 new cases of ovarian cancer each year, accounting for around 4% of all cancers diagnosed in women (IARC, 2004). Incidence rates vary considerably, with the highest rates in the USA and Northern Europe and the lowest rates in Africa and Asia (IARC, 2004). Around 43,000 cases occur each year in Europe and 22,000 in the USA. Within the EU the lowest rates are in the Southern European countries of Greece, Portugal and Cyprus, while the highest are in the Northern and Eastern European countries of Lithuania, Denmark, Czech Republic and Estonia (IARC, 2004).
Evidence from population migrant studies is conflicting. In a study of immigrants to Israel, those who migrated from a country with a higher incidence of ovarian cancer retained a similar level of risk as their country of origin, but within one generation the risks were no longer significantly different (Parkin and Iscovich, 1997). However, research in other migrant populations has not shown a similar pattern (Hamminki and Li, 2002a; Hamminki and Li, 2002b; Hamminki et al., 2002; Harding, 1998).

The most important prognostic factor in ovarian cancer remains the stage of the disease at presentation. Unfortunately, around 75% of affected women present with advanced stage (FIGO III, IV) disease.

Ovarian cancer has been described as a ‘silent killer’ owing to the fact that most women with this disease remain asymptomatic in the early stages. Granberg and Wikland (1988) found in a retrospective study that the most common presenting complaint in women with FIGO Stage I and II was abdominal discomfort (64%). 10% of women were completely asymptomatic, and 8% presented with postmenopausal bleeding.

A recent retrospective study by Lataifeh et al. (2005), which included 100 women with early stage ovarian cancer and 100 women with advanced stage ovarian cancer, showed that 90% of women with early stages ovarian cancer and 100% of women with advanced ovarian cancer had at least one symptom. 51% of women with early disease reported abdominal pain and 32% reported abdominal distension. However, 62% of women with advanced disease reported abdominal swelling and 44% reported abdominal pain. Seventy percent of the early stage and 69% of the advanced stage cohorts reported symptoms of less than 3 months duration. Tumours less than 5 cm diameter were three times more likely to have advanced disease (p=0.02). The authors concluded that advanced disease is not invariably due to delayed diagnosis (Lataifeh et al., 2005).
Women with early stage ovarian cancer are more likely to report an abdominal mass or urinary symptoms but less likely to report gastrointestinal problems or general malaise compared to advanced stage cancer. General malaise and other non-specific symptoms are least common in borderline disease. Older women, and those with higher parity or a family history of breast or ovarian cancer, are more likely to be diagnosed at an advanced stage of disease (Webb et al., 2004).

Most malignant ovarian tumours are epithelial ovarian cystadenocarcinomas. Tumours of low malignant potential comprise approximately 20% of malignant ovarian tumours, fewer than 5% of malignant germ cell tumours, and approximately 2% of granulosa cell tumours.

Different risk factors have been associated with ovarian cancer. Examples of risk factors include exposure to an increased number of ovulatory cycles like with early menarche, late menopause, nulliparity and recurrent fertility treatment. The use of the oral contraceptive pill appears to have a protective effect against ovarian cancer (Edmondson and Monaghan, 2001). Genetic predisposition like the one seen in women with BRCA1 and BRCA2 gene mutations (Malander et al., 2004), and family history of ovarian, breast or colorectal cancer are also risk factors for ovarian cancer.
Figure 1 Number of new cases and age specific incidence rates, female ovarian cancer, UK 2004 (Cancer Research UK, http://info.cancerresearchuk.org/cancerstats/faqs/?a=5441,11/2007)
5.3 Natural History of Ovarian Cancer

The natural history of ovarian carcinoma remains unknown. A study by Horiuchi et al. (2003) suggested that approximately half of ovarian carcinomas develop from pre-existing, benign appearing cysts and they called this pathway adenoma-carcinoma sequence, whereas the remaining half seem to develop from a normal appearing ovary and they called this pathway de novo carcinogenesis. This study suffers from being a retrospective study containing a selected mix of only forty-nine women.

Another proposed theoretical model by Shih and Kurman (2004) suggested that surface epithelial tumours are divided into two broad categories designated type I and type II tumours, which correspond to two main pathways of tumourigenesis. Type I tumours tend to be low-grade neoplasms that arise in a stepwise manner from borderline ovarian tumours. This group includes mucinous carcinomas, endometrioid carcinomas, malignant Brenner tumours, and clear cell carcinomas. Type I tumours are associated with distinct molecular changes such as BRAF and KRAS mutations for serous tumours, KRAS mutations for mucinous tumours, and β-catenin and PTEN mutations and microsatellite instability for endometrioid tumours.

Type II tumours are high-grade neoplasms, which arise from normal ovaries (de novo development). This group of tumours includes high-grade serous carcinoma, malignant mixed mesodermal tumours (carcinosarcoma), and undifferentiated carcinoma. There are very limited data on the molecular alterations associated with type II tumours except frequent p53 mutations in high-grade serous carcinomas and malignant mixed mesodermal tumours.

Shih and Kurman (2004) suggested that their model for carcinogenesis reconciles the relationship of borderline ovarian tumours to invasive carcinoma and provides a morphological and molecular framework for studies aimed at elucidating the pathogenesis of ovarian cancer.
This study is theoretical and there is no clinical proof as yet to support these theories of tumourigenesis.

5.4 Benign Ovarian Tumours

Benign ovarian tumours have increased in prevalence in the last few years due to the increased use of ultrasound scanning in our everyday clinical practice (Valentin, 2004). Many adnexal masses that would almost certainly have remained undetected before the ultrasound era are now found incidentally at transvaginal ultrasound examination of women without symptoms of an adnexal tumour.

In modern clinical practice it is not acceptable to operate on all asymptomatic women with an incidental ultrasound diagnosis of a benign looking cyst. Some of these cysts may be functional in nature and would resolve spontaneously, while others may need to be operated on.

Laparoscopic treatment of ovarian cysts has been established as the optimal route for surgical management of benign adnexal tumours (Mahdavi et al., 2004). This minimally invasive approach offers many advantages over laparotomy, including decreased postoperative pain, faster recovery and reduction in postoperative intra-abdominal adhesions (Lin et al., 1995; Howard, 1995; Mais et al., 1995). As a result, many operations can be performed as day-surgery procedures, with patients being discharged home few hours after their operation.

Certain ovarian cysts may manifest typical appearance on ultrasound and therefore could be diagnosed with confidence (Patten, 1992; Guerriero et al., 1995; Kim et al., 1995, Kupfer et al., 1992; Caspi et al., 1996). Other tumours may have features similar to malignant tumours and these could be misdiagnosed and over-treated as a result.

Detailed below, are the different benign ovarian tumours that can occur.
5.4.1 Follicular cysts

These cysts arise from the ovarian follicles and may occur at any age before the menopause and could reach large sizes. Their lining cells (granulosa and theca cells) may secrete excessive quantities of oestrogens.

On ultrasound, these cysts appear as thin-walled cysts with anechoic fluid content and they are surrounded by healthy ovarian tissue. Occasionally, haemorrhage into the cyst may result in a web-like appearance.

The vast majority of these cysts are asymptomatic and tend to resolve spontaneously. However, in order to be able to reach a diagnosis of follicular cyst, a follow up ultrasound is usually recommended and this should be performed usually in the first week of the woman's menstrual cycle.

5.4.2 Corpus Luteum cysts

The corpus luteum is a direct result of ovulation from a follicle. The morphology of the corpus luteum vary throughout the luteal phase, and will be predominantly cystic or solid in nature.

On ultrasound, the corpus luteum tends to have a haemorrhagic cystic appearance containing spider-web-like material (Grant, 1992; Patel et al., 1999). Odd looking blood clots may also be seen (Carter et al., 1992). These clots may be confused with papillary projections or solid components even by the best trained eye and this may lead to them being misdiagnosed as malignant lesions. Doppler ultrasound examination may help to discriminate between a clot and a solid component, which tend to have detectable Doppler signals. The blood clot will move like jelly when the ovary is pushed using the vaginal probe.
Again, expectant management is the option of choice in the case of corpus luteum cysts and a follow up with ultrasound after 6–12 weeks may be offered, if in doubt about the diagnosis.

5.4.3 Theca-Lutein cysts

These cysts tend to occur when a corpus luteum becomes cystic and persist in a functional state for longer than normal. Alternatively these cysts occur when the granulosa and theca cells of a follicular cyst become luteinised.

On ultrasound, these cysts have a unilocular smooth-walled appearance with anechoic content.

The majority of these cysts resolve spontaneously and expectant management is recommended.

5.4.4 Serous inclusion cysts

These cysts form as a result of cortical invagination of the surface germinal epithelium. They are typically less than 3 cm in diameter. On ultrasound, they appear as unilocular, smooth-walled cysts with anechoic content. Expectant management is also recommended here.
5.4.5 Endometriomas

The ovarian endometrioma represents an advanced stage of endometriosis. Although many reports describe them as having typical ultrasound features, they are a common source of false positive diagnosis of malignancy. They are typically seen on ultrasound examination as well circumscribed thick walled cysts that contain homogeneous low level internal echoes. This so-called ‘ground glass’ appearance is due to altered blood. The fluid is often hypoechoic so in some cases it may be necessary to increase the gain setting to detect the low-level echogenicity (Tailor et al., 1999). Internal septations have been described in 10-30% of all endometriotic cysts seen, which may cause diagnostic difficulties (Athey and Diment, 1989; Kupfer et al., 1992; Volpi et al., 1995). Patel and co-workers (1999) noted ‘solid masses’ protruding from the endometriotic cyst wall into the cyst lumen. Such protrusions (‘wall nodularity’) are seen in 20% of endometriomas (Patel et al., 1999). They reported that ‘hyperechoic wall foci’ were much more common in endometriomas than in other types of adnexal pathology (35% versus 6%), a finding that has not been reported by others.

In addition, ovarian endometriomas may undergo decidualisation in pregnancy, which sometimes creates confusing morphological and Doppler features (Fruscella et al., 2004).

While operative laparoscopy is the gold standard procedure to treat endometrioma, there was no proven role for the medical management of these cysts. The use of cyclical oral contraceptive pill postoperatively does not seem to affect the long-term recurrence of endometriosis after surgical removal (Chapron et al., 2002).
5.4.6 Epithelial tumours

These tumours arise from the ovarian surface epithelium and the majority of benign epithelial tumours could be either serous or mucinous.

Benign serous ovarian cystadenomas are the commonest in benign epithelial tumours and they are usually seen in women over the age of 40 (Girling and Soutter, 1997). 10% of serous cystadenomas are bilateral. The characteristic ultrasound features of serous cystadenomas are unilocular or bilocular cysts with homogeneous echogenicity (often anechoic) with a thin regular wall and occasionally with a thin regular septum and no papillary projections (Buy et al., 1991). Although, Russell in 1979 found that papillary projections were identified in 4.5% of benign serous cystadenomas (Russell, 1979). The prevalence of papillary projection in this type of benign ovarian tumour is likely to be higher as demonstrated later in the results of this thesis. Valentin et al. (2001) showed no statistically significant difference in the prevalence of papillary projections between benign and malignant ovarian tumours. Papillary projections were present in up to 26% of benign tumours compared to 28% of malignant ovarian tumours (p=0.88). The presence of irregular thickenings along the septa of a multilocular thin-walled cyst may be representative of a serous cystadenofibroma.

Mucinous cystadenomas account for approximately 15-20% of all ovarian tumours. 10% of these are bilateral and tend to be larger in size compared to serous cystadenomas. Mucinous cystadenomas could be unilocular in 35% of the cases; however the majority of these cysts are multilocular. On ultrasound, these tumours appear as multilocular cysts containing fluid of different echogenicities, with regular wall and septa (Buy et al., 1991). They may also exhibit a unilocular smooth-walled cyst appearance with echogenic fluid content. Papillary projections may occasionally be noted in the cysts.
Other less common types of benign epithelial ovarian tumours include Brenner (transitional cell) tumours, mixed epithelial and endometrioid. Brenner tumours tend to be solid, well circumscribed, hypoechoic tumours with calcification and acoustic shadowing on ultrasound examination.

Buy et al. (1991) diagnosed serous cystadenomas with a sensitivity of 70%, specificity of 98%, LR+ of 35 and a LR- of 0.31, and mucinous cystadenomas with a sensitivity of 50%, specificity of 96%, LR+ of 12.5 and a LR- of 0.52 using the ultrasound features described above. Fleisher and co-workers (1978) found the sensitivity of gray-scale imaging for diagnosing serous cystadenoma to be 75%, the specificity to be 75%, LR+ to be 3 and LR- to be 0.33. The sensitivity, specificity, LR+ and LR- of gray-scale imaging for diagnosing mucinous cystadenomas were reported by Fleischer and co-workers to be 95%, 99%, 31.67 and 0.05 respectively (Fleisher et al., 1987).

These tumours could be managed expectantly, if the patient is asymptomatic, with regular follow up scans. However, the gold standard management is by laparoscopic surgical excision of the cysts.

5.4.7 Germ cell tumours

Mature cystic teratomas (Dermoid cysts) are the most common type of germ cell tumours. They develop from totipotential cells and consist of well differentiated ectodermal and mesodermal elements. Approximately 10-12% of all ovarian neoplasms are benign cystic teratomas, which are bilateral in 10-15% of patients.

Most dermoid cysts are easily recognised at gray-scale imaging owing to their fat and hair content (Caspi et al., 1996; Cohen and Sabbagha, 1993). Having said that, around 18% of dermoid cysts may manifest a predominantly cystic echo pattern indistinguishable from that of other cystic masses (Cohen and Sabbagha, 1993). The
most characteristic ultrasound features of a dermoid cyst are the presence of (1) a ‘white ball’ (corresponding to hair and sebum) in the corner of a cyst, or filling up the whole tumour; (2) long, echogenic (white) lines and prominent echogenic dots in cyst fluid (corresponding to hair floating freely in non-fatty fluid); and (3) shadowing (Caspi et al., 1996; Cohen and Sabbagha, 1993; Bronshtein et al., 1991). The latter often makes it difficult or impossible to measure the size of dermoid cysts correctly. Some dermoid cysts are impossible to detect, even when they are clearly palpable, because their echogenicity is similar to that of surrounding bowel (Valentin, 2004).

Dermoid cysts are managed by laparoscopic ovarian cystectomy. The majority are less than 10 cm, and 15% are bilateral. If ruptured intra-operatively, copious irrigation is warranted because the content can cause a chemical peritonitis (Huus et al., 1996; Coccia et al., 1996; Remorgida et al., 1998). Therefore some authors advocate open surgery for large dermoid cysts >10 cm in mean diameter as they are difficult to remove from the abdominal cavity without rupture (Mettler et al., 2001).

5.4.8 Sex cord-stromal tumours

The stromal cells retain the potential for differentiation into any of the cells or tissues arising from the mesenchyme of the gonad. These tumours, which are also known as gonadal stromal tumours, mesenchymomas, may secrete any or all of the ovarian steroids. Thecal cell tumours consist of:

- Fibromas: consist mainly of fibroblast-like cells producing collagen. These are large, firm and mobile tumours, which are bilateral in 10% of cases. Ascites may be present in 15% of cases. They could be part of a complex known as Meig’s syndrome, which is a triad of a benign solid ovarian tumour, ascites and pleural effusion (usually right sided) (Meigs and Cass, 1937). Both ascites and pleural effusion tend to resolve after removal of the tumour (Nemeth and Patel, 2003).
Thecomas: predominantly contain lipid-rich theca-like cells

Fibromas and fibrothecomas are solid, round or oval tumours with a smooth outline and a regular 'striped' echogenicity and usually have an echo pattern indistinguishable from that of pedunculated uterine fibroids. Different types of solid benign ovarian tumours, such as fibroma, thecoma, fibrothecoma and Brenner tumour may have similar appearance and echogenicity on gray-scale ultrasound (Athey and Malone, 1987). Guerriero and co-workers (1995) reported confusion of ovarian endometrioma and fibroma at gray-scale sonography.

The management of these tumours depend on their size. Furthermore, the presence of a solid tumour may raise suspicion of malignancy and additional investigations such as serum tumour markers may be used. The choice between laparoscopy and laparotomy will depend on the operating surgeon and the facilities available to remove the solid lesion via a small abdominal incision.

5.4.9 Other Benign Adnexal Tumours

Lesions adjacent to the ovary occasionally present in a similar fashion to ovarian tumours. These are not infrequently seen especially in young women and in those who have undergone previous pelvic surgery.

5.4.9.1 hydro-pyo-haemato-salpinx

These appear as fluid-filled tubular cystic structures, which sometimes are easily distinguishable from other types of adnexal masses at gray-scale ultrasound examination. Characteristic ultrasound features of tubal disease include: fluid-filled sausage-shaped cystic structure, presence of 'incomplete septa', i.e. septa that are not seen to reach the opposite wall of the cystic structure, and on a transverse section of these lesion, mucosal folds are seen to protrude into the lumen, resulting in a 'cog-
wheel' appearance if the tube is swollen and in a 'beads-on-a-string' appearance (Timor-Tritsch and Lerner, 1998).

Management of tubal disease is usually conservative with analgesia and antibiotics, like in the case of a pyosalpinx. Surgery may be carried out on infertility patients and this will be in the form of bilateral salpingectomy in order to improve the success of in vitro fertilisation (IVF) treatment.

### 5.4.9.2 Tubo-ovarian abscess

The ultrasound appearance of a tubo-ovarian abscess may vary from a unilocular cystic structure to a complex multicystic lesion containing thick walls and thick septa, filled with homogeneously echogenic material ('ground-glass' appearance) (Valentin, 2004). These abscesses may be confused with endometriomas or ovarian malignancies (Hata et al., 1989; Varras et al., 2003). Differentiating a tubo-ovarian abscess from a pelvic abscess of other origin such as peri-appendiceal or diverticular abscess is not an easy task (Valentin, 2004). Thus, a full clinical history and results of blood tests are important in order to establish the correct diagnosis.

Tubo-ovarian abscesses are usually treated conservatively with broad spectrum antibiotics and occasionally by ultrasound guided transvaginal drainage.

### 5.4.9.3 Paraovarian cysts

These cysts may arise from embryonic ducts and may be mesothelial, mesonephric and paramesonephric in origin (Genadry et al., 1977). They are usually situated between the tube and the ovary (Grant, 1992). These cysts occur most commonly in premenopausal women and account for approximately 10% of all adnexal lesions (Athey and Cooper, 1985).
Paraovarian cysts may appear as cysts, which are clearly separate from a normal ovary on ultrasound examination; on other occasions it is difficult to distinguish an ovarian cyst from a paraovarian cyst. They usually appear as unilocular anechoic cysts but may have echogenic fluid content occasionally. Papillary projections and septa may be present (Korbin et al., 1998). Malignancy, especially in the form of a borderline tumour rather than an invasive malignancy may develop in a paraovarian cyst (Stein et al., 1990).

These cysts can be managed expectantly with regular follow up scanning. However, transvaginal drainage or surgical removal may be indicated if the patient becomes symptomatic or the cyst changes in appearance.

5.4.9.4 Peritoneal pseudocysts
Peritoneal pseudocysts are fluid collections among adhesions occurring after an inflammatory process in the peritoneal cavity or after an operation. On ultrasound examination, peritoneal pseudocysts appear as cystic lesions following the contours of the pelvis with a deformed ovary suspended among adhesions centrally or peripherally in the cyst. Occasionally, pseudocysts may also be oval or round with anechoic or echogenic fluid content. Thin septa may be noted within the cysts and these septa appear to move in response to pressure from the endovaginal probe (flapping sail sign) (Savelli et al., 2004). Papillary projections may sometimes be noted within pseudocysts (Jain, 2000; Hoffer et al., 1988; Kim et al., 1997).

5.5 Borderline Ovarian Tumours (BOTs)
BOTs account for 10-15% of all ovarian tumours. They commonly affect white women in the fourth decade of life, but they are seen in women of all ages (Silverberg et al., 2004). They are usually localised to the ovary at the time of the initial presentation
The reported annual incidence of serous and mucinous BOTs is 24.7 per one million women, the incidence increasing with increasing age but at a slower pace than for invasive malignant ovarian tumours.

Serous and mucinous BOTs constitute around 96% of BOTs. The remaining types are clear cell, endometrioid, Brenner (transitional cell) or mixed histology. Serous BOTs are subdivided into typical serous BOTs and micropapillary/cribriform serous BOTs (Seidman and Kurman, 1996). The majority of serous BOTs are of typical subtype (Gilks et al., 2003). This type of BOTs is bilateral in 35-40% of the cases and is similar in gross histological appearance to benign serous ovarian tumours. However, papillary excrescences are more common and likely to have surface component (70% of lesions) (Russell et al., 2002a). Micropapillary serous BOTs have higher frequency of bilaterality (Ahmed and Lawton, 2005).

Mucinous BOTs are subdivided into gastrointestinal (GI) type, which is the most common type, and endocervical (Mullerian or seromucinous) type. GI type mucinous BOT is usually unilateral, whilst endocervical type mucinous BOT is bilateral. On macroscopic histological examination, GI type mucinous BOTs are large tumours reaching an average of 15-20 cm in mean diameter. The external surface of the tumour and its cyst walls are generally smooth. These tumours have tendency to a finer honeycombing of the cut surface in higher-grade proliferating tumours but this may be only focal (Russell et al., 2002b). These tumours are tensely filled with mucin. Endocervical mucinous BOTs are characterised by the presence of prominent intracystic and exophytic papillary structures resembling serous BOTs (Hart, 2005).

Several recent studies have examined the gray-scale sonographic and colour Doppler features of borderline ovarian tumours (Exacoustos et al., 2005; Pascual et al., 2002; Wu et al., 1994). There were also attempts to describe the typical sonographic
features of different histopathological subtypes of BOTs (Fruscella et al., 2005; Exacoustos et al., 2005; Gotlieb et al., 2000).

Fruscella et al. (2005) and Exacoustos et al. (2005) retrospectively reviewed the ultrasound appearance of BOTs. They reached a conclusion that serous BOTs and endocervical type mucinous BOTs have a common ultrasonic appearance of a unilocular cyst with papillary projections arising from the inner wall. They also found that GI type mucinous BOTs have the appearance of a multilocular (containing more than 10 locules on most occasions) cyst with mucinous fluid or a cyst. These appearances coincided with the gross histopathological description of these tumours.

However, the diagnostic accuracy of pattern recognition for the differential diagnosis between benign, borderline and invasive ovarian tumours has not been tested as yet.

Most BOTs have a good prognosis. Important prognostic factors are the histological subtype and tumour stage. Accurate surgical pathological staging is important (Ahmed and Lawton, 2005). The 5-year survival is nearly 100% for serous BOTs in stage I (Zanetta et al., 2001). In cases of extraovarian disease, 5-year survival is more than 85% if the implants are non-invasive and around 60% if the implants are invasive (exhibiting irregular infiltration of the surrounding or underlying normal tissue, and usually resembling low-grade serous carcinoma histologically). However, lesions demonstrating these features are quite rare in most series (Silverberg et al., 2004). Micropapillary type serous BOTs are more likely to be diagnosed at an advanced stage, and their prognosis is worse than that of typical serous BOTs (Ahmed and Lawton, 2005). Almost all mucinous BOTs have an excellent prognosis with a 5-year survival rate approaching 100% after surgical treatment (Silverberg et al., 2004). In the past, pseudomyxoma peritonei, which is associated with a high mortality rate, used to be classified as mucinous BOT. However, pseudomyxoma peritonei is now believed to be
of gastrointestinal (often appendix) - not ovarian - origin and is therefore no longer classified as mucinous BOT (Ronnett et al., 2004).

Management is surgical and is divided into conservative (unilateral oophorectomy) or radical (total hysterectomy, bilateral salpingo-oophorectomy and omental biopsy), depending on the patient's age and reproductive potential.

5.6 Malignant ovarian tumours
The appearance and behaviour of invasive ovarian malignancies depend on the histological type of these tumours.

5.6.1 Epithelial ovarian cancers
Epithelial ovarian cancers (cystadenocarcinoma) represent 85% of all primary malignant neoplasms. These tumours affect women aged 50-70 years. This type of ovarian cancer is associated with raised serum CA-125 in over 85% of the cases. Bilateral tumours are noted in 50% of serous cystadenocarcinomas and in 25% of mucinous cystadenocarcinomas.

On ultrasound examination, serous cystadenocarcinomas may be either unilocular or multilocular and are associated with varying proportions of a solid component (Troiano and McCarthy, 1994; Yamashika et al., 1995). The presence of papillary projection is typical in early stage cystadenocarcinomas, which makes them difficult to differentiate from serous borderline ovarian tumours (Valentin et al., 2006a). In multilocular tumours, the septa are often thick and nodular. These cysts are large and in around half of the cases they are greater than 15 cm in diameter. Doppler examination shows increased vascularity to the tumour and 50% of cases has ascites present.
Mucinous cystadenocarcinomas are larger than their serous counterpart. They appear as multilocular cysts containing echogenic fluid material and extensive papillary projections on ultrasound examination. These tumours also show high vascularity on Doppler examination. Mucinous (gelatinous) ascites may be found and this is difficult to drain by paracentesis, thus patients may have to undergo laparotomy for the drainage purpose.

Endometrioid ovarian cancers represent 10-25% of all epithelial ovarian cancers. Endometrial adenocarcinomas coexist with up to 30% of endometrioid ovarian cancer (Cotran et al., 1994). 30-50% of cases are bilateral and they appear as cystic masses containing extensive papillary projections on ultrasound examination. These tumours are also characterised by increased vascularity on Doppler examination.

5.6.2 Sex-cord stromal ovarian cancers

Six percent of ovarian cancers are sex-cord stromal cancers. These tumours arise from the stromal cells and thus secrete any or all of the ovarian steroids. Granulosa cell tumours are the commonest subtype. Five percent of these tumours happen before puberty and the remainder over the reproductive and postmenopausal years.

The appearance of these tumours varies considerably on ultrasound examination. Large tumours have a complex appearance and could be mistaken for epithelial cancers, whereas smaller tumours may appear as solid hypoechoic masses similar to ovarian fibromas. These tumours have high vascularity on Doppler examination.

Because of their hormonal secreting characteristics, these tumours may present with postmenopausal bleeding owing to endometrial hyperplasia or endometrial carcinoma in view of the oestrogen secreted by this type of ovarian tumours.
Granulosa cell tumours present in early stages, they have a low malignant potential, and subsequently a lower mortality rate and better prognosis compared to epithelial ovarian cancers.

Other less common subtypes of sex-cord ovarian cancers include Sertoli-Leydig cell tumours, which may secrete virilising hormones and often appear as solid masses on ultrasound examination.

5.6.3 Germ cell ovarian cancers

Three percent of ovarian cancers fall into this category. These cancers tend to occur more frequently in younger women between 20-30 years of age. Histological subtypes include; choriocarcinomas, yolk sac tumours and dysgerminomas. The dysgerminomas are the commonest subtype.

On ultrasound examination, these tumours appear as solid and hyperechoic masses.

Germ cell tumours can produce pregnancy associated plasma proteins such as human chorionic gonadotrophin or alpha-fetoprotein, which may be useful as tumour markers in the diagnosis of this type of ovarian cancer and its subsequent follow-up after treatment.
5.6.4 Metastatic ovarian tumours

The ovary is a common site of metastases from malignant tumours, 5–20% of ovarian masses being metastases from primary tumours in other organs (Young and Scully, 1991). Most (50–90%) metastases in the ovaries originate from the gastrointestinal tract or the breast (Antila et al., 2006).

Metastatic tumours are described as often (about 60% of cases) being bilateral lesions that appear as diffusely solid tumours, multiple solid nodules, partly cystic masses or, rather uncommonly, entirely cystic lesions. Even when the primary tumour is solid, metastases in the ovaries may be cystic or semicystic. Extensive areas of haemorrhage and/or necrosis are common. Krukenberg tumours (metastasis from a gastric cancer to the ovary) are typically solid masses with a bosselated outer surface.

Testa et al. (2007) retrospectively reviewed the ultrasound appearance of 67 metastatic tumours in the ovary. They found that ovarian metastases derived from stomach cancer, breast cancer, lymphomas and uterine cancer are solid in almost all cases, whereas those derived from the colon, rectum and biliary tract manifest more heterogeneous morphological patterns, most being cystic with many cyst locules and irregular borders. Papillary projections seem to be relatively rare (12%) in metastatic tumours in the ovaries.

5.7 Management of Ovarian Cancer

Management of ovarian cancer is usually guided by the histological grade and the stage of the tumour. The best survival rates for women with ovarian cancer are achieved when treatment is organised and carried out by gynaecological oncologists working in cancer centres (Junor et al., 1999). These cancer centres have been recently introduced across the UK. Ideally all patients with suspected or proven ovarian cancers should be referred to these tertiary centres for management (Kehoe et al., 1994).
The treatment of choice for stage I ovarian cancer is surgical staging with laparotomy. Although simple oophorectomy at the time of surgical staging may be considered in younger women with stage I disease who wish to preserve their reproductive potential. Cytoreductive surgery is advocated for patients with more advanced stages of the disease, followed by chemotherapy.

The International Federation of Gynecology and Obstetrics (FIGO) has produced strict criteria for the accurate staging of ovarian cancers (Table 1) (Shepherd, 1988).

The 5-year survival rates vary and depend on the stage and the histological type of the tumour. Stage Ia epithelial cancers have a 5-year survival rate of 70%, whilst with stage IV disease the survival rate falls to below 5%. The survival rates for gonadal-stromal and germ cell tumours are better stage for stage than those for epithelial cancers.
Table 1 International Federation of Gynecology and Obstetrics (FIGO) staging of primary ovarian carcinoma.

<table>
<thead>
<tr>
<th>FIGO Stage</th>
<th>Tumour characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Growth limited to the ovaries</td>
</tr>
<tr>
<td>Ia</td>
<td>Growth limited to one ovary, no ascites present containing malignant cells, no tumour on the external surface, capsule intact</td>
</tr>
<tr>
<td>Ib</td>
<td>Growth seen in both ovaries, no ascites present containing malignant cells, no tumour on the external surfaces, capsules intact</td>
</tr>
<tr>
<td>Ic</td>
<td>Tumour either stage Ia or Ib but tumour found on surface of one or both ovaries, capsule ruptured, ascites present containing malignant cells, or peritoneal washings positive</td>
</tr>
<tr>
<td>II</td>
<td>Growth involving one or both ovaries with pelvic extension</td>
</tr>
<tr>
<td>IIa</td>
<td>Extension or metastases to the uterus or fallopian tubes</td>
</tr>
<tr>
<td>IIb</td>
<td>Extension to other pelvic tissues</td>
</tr>
<tr>
<td>IIc</td>
<td>Tumour either stage IIa or IIb but tumour found on surface of one or both ovaries, capsule(s) ruptured, ascites present containing malignant cells, or peritoneal washings positive</td>
</tr>
<tr>
<td>III*</td>
<td>Tumour involving one or both ovaries with histologically confirmed peritoneal implants outside the pelvis or positive retroperitoneal or inguinal nodes; tumour limited to the true pelvis but with histologically confirmed malignant extension to small bowel or omentum</td>
</tr>
<tr>
<td>IIIa</td>
<td>Tumour grossly limited to the true pelvis with negative nodes but with histologically confirmed microscopic seeding of abdominal peritoneal surfaces or extension to small bowel or mesentery</td>
</tr>
<tr>
<td>IIIb</td>
<td>Tumour involving one or both ovaries with histologically confirmed implants, metastasis to abdominal peritoneal surfaces (lesions ≤2 cm), negative nodes</td>
</tr>
<tr>
<td>IIIc</td>
<td>Extrapelvic peritoneal metastasis (lesions &gt;2 cm) or positive retroperitoneal or inguinal nodes</td>
</tr>
<tr>
<td>IV†</td>
<td>Growth involving one or both ovaries with distant metastases, positive cytologic findings if pleural effusion is present</td>
</tr>
</tbody>
</table>

Source: Shepherd 1988

*Includes superficial liver metastases.
†Includes parenchymal liver metastases.
PART TWO:

Aims, Methods and Results
Chapter 6: Aims

6.1 Aims of the thesis

- Compare the value of the ovarian crescent sign and that of the risk of malignancy index for the diagnosis of ovarian cancer in women with an ultrasound diagnosis of ovarian tumour.

- Assess the value of pattern recognition for the differential diagnosis of adnexal tumours and in particular its accuracy in establishing the specific diagnosis of histological subtypes of borderline ovarian tumours.

- Determine if the prediction of the character of an adnexal mass using subjective evaluation of gray-scale and Doppler ultrasound findings (i.e. pattern recognition) based on the evaluation of static images is as accurate as that based on a real-time ultrasound examination.

- Evaluate the reproducibility of ultrasound 'pattern recognition' for the diagnosis of borderline ovarian tumours.

- Examine the degree of confidence with which expert ultrasound operators differentiate between benign, borderline and invasive malignant ovarian tumours and assess the effect of their diagnostic confidence on diagnostic accuracy and interobserver agreement.

- Examine whether the improved accuracy of expert (level III) ultrasonography has a measurable effect on the management of patients with adnexal masses when compared with routine (level II) ultrasonography.

- Develop and assess a set of ultrasound criteria that were designed to select women for laparoscopic removal of ovarian cyst with a minimum risk of converting the operation to a laparotomy.
Chapter 7: Materials, Methods and Results

7.1 Setting
The Early Pregnancy and Gynaecology Assessment Unit, King’s College Hospital, London.

7.2 Gynaecology population
The work contained at this thesis was carried out at the Early Pregnancy and Gynaecology Assessment Unit, King’s College Hospital between April 2005 and October 2007.

The hospital is a teaching hospital located in South East London and is one of the main two teaching hospitals for Guy's, King's and St Thomas' Schools of Medicine and Dentistry. It serves the health needs of Lambeth, Southwark and Lewisham Health Authority. The local population of approximately 750,000 people is one of the most deprived in England with high levels of premature death, mental illness and infectious diseases including HIV.

King’s College Hospital became a NHS Foundation Trust in December 2006. It is one of London’s largest and busiest teaching hospitals, with a unique profile of strong local services and a focused set of tertiary specialties. It has 900 beds and employs 4000 staff with an annual budget of approximately £185 million. The women’s health directorate makes up one of the largest departments and offers specialist services including gynaecological ultrasound scanning, gynaecological oncology, urogynaecology, fetal medicine, colposcopy, menopause, assisted conception and family planning. In 2006, a total of 17000 women were referred to the Early Pregnancy and Gynaecology Assessment Unit, half of whom were seen for variety of gynaecological disorders. A significant portion of these women were referred from
other hospitals in the UK with adnexal lesions or difficult early pregnancy complications.

All the women in the randomised controlled trial were recruited from the Rapid Access Unit of the Regional Gynaecological Cancer Centre, Guy’s and ST Thomas’ NHS Foundation Trust. A proportion of women included in the remaining studies referred from the Regional Gynaecological Cancer Centre or the Obstetrics and Gynaecology department at Guy’s and ST Thomas’ NHS Foundation Trust. This hospital is the other large teaching hospital for Guy’s, King’s and St Thomas’ Schools of Medicine and Dentistry. This trust, which was given a Foundation status in 2004 employs around 9,000 staff and provides three-quarter million patient contacts in acute and specialist hospital services every year. The Gynaecological Cancer Centre is one of the main tertiary gynaecology cancer centers in the South East.

7.3 Preoperative assessments
Preoperative assessment took place either at King’s College Hospital or at Guy’s and St Thomas NHS Foundation Trust. Women with known adnexal pathology were recruited for the different studies where an informed consent was obtained from all patients.

Women were examined preoperatively using transvaginal gray-scale and colour Doppler ultrasound in the dorsal lithotomy position. Age and menopausal status were recorded in all women. Postmenopausal status was defined as more than one year of amenorrhoea or age more than 50 years in women who had had a hysterectomy.

7.4 Randomisation
Randomisation of patients into the randomised controlled study took place at the Regional Gynaecological Cancer Centre, Guy’s and ST Thomas’ NHS Foundation Trust. An information leaflet about the study was given to all eligible patients before
assessment. The patients were reviewed by a member of the gynaecological oncology team who took a detailed clinical history, did a physical examination, and arranged additional tests when needed. If the diagnosis of an adnexal mass was confirmed, ultrasonography was requested. Written informed consent was obtained from all patients who agreed to take part in the study.

Consecutively numbered, opaque sealed envelopes were prepared at King's College Hospital NHS Foundation Trust using a computer-generated randomisation sequence. The envelopes were then taken to the gynaecological rapid access clinic at Guy's and St Thomas’ NHS Foundation Trust and securely kept in the nurses' office. The envelopes were distributed by the nursing staff to the clinicians who were responsible for randomisation. The patients were randomised into either having a level II or a level III ultrasonography and each envelope contained the name of the department to which the patient was allocated.

7.5 Ultrasonography

Ultrasound examinations were performed using an Aloka SSD-5000 or an Aloka SSD-2000 machine (Aloka Co, Tokyo, Japan) or a Siemens Antares™ machine (Siemens Medical Solutions, UK). Transvaginal scan was performed except in cases were it was contraindicated, such as virgins or those unable to consent. Transabdominal sonography was used to examine a large mass that could not be seen in its entirety using a transvaginal probe or where transvaginal ultrasonography was not possible.

Transvaginal ultrasound examination was performed using a 5 MHZ transducer with B-mode and Doppler facilities, incorporating a filed view of 90°. The following morphological ultrasound information was recorded in each case: volume of cyst (s), presence of septations and solid areas within the cyst, metastases and ascites. An attempt was made at identifying healthy ovarian tissue adjacent to the tumour (the
ovarian crescent sign). This was defined as visible hypoechogenic tissue with or without ovarian follicles enclosed within the ovarian capsule and located adjacent to the cyst wall. This tissue would not be separated from the cyst when applying a moderate amount of pressure. Measurements of the diameter of each cyst were taken in three perpendicular planes. The tumour or solid part mean diameter was calculated according to the formula \((D_1 + D_2 + D_3)/3\) and the tumour volume was calculated according to the formula \((\pi/6 \times D_1 \times D_2 \times D_3)\). Transabdominal scans were performed in women with large pelvis masses extending out of the range of the vaginal probe to ensure a complete examination or in cases where the patient was a virgin.

The nature of the ovarian cyst was assessed by use of the pattern recognition method. Features considered suspicious of an ovarian cancer included: presence of ascites and peritoneal tumour deposits within the pelvis; extensive papillarities (defined as the presence of multiple papillary (>2) projections >3mm in height covering large sections of the inner cyst wall) arising from the inner cyst wall (Valentin, 2004; Timmerman et al., 2000; Jokubkiene et al., 2007); and irregular solid components within the cyst with evidence of necrosis (i.e. ill-defined anechoic areas within a predominantly solid lesion). Metastases were defined as the presence of irregular solid tumour deposits within the pelvis or the presence of omental cake. Operators were asked to classify tumours as either benign or malignant. When such estimates could not be made, the ultrasonography findings were classified as non-diagnostic and only a morphological description of the tumour was provided to the clinicians.

Level III ultrasonography was done by gynaecologists who had a special interest in gynaecological ultrasonography; had more than 10 years’ experience in gynaecological ultrasonography; worked in a tertiary referral unit; had been continuously involved in research in the field of gynaecological ultrasonography; and were recognised preceptors for the training of gynaecological ultrasonography by the
UK Royal College of Obstetricians and Gynaecologists. Thus, they satisfied the criteria for level III operators as specified by the European Federation of Societies for Ultrasonography in Medicine and Biology (EFSUMB Newsletter, 2005).

Level II ultrasonography was done by ultrasonographers who were trained in gynaecological ultrasonography and all routinely and independently assessed patients with adnexal tumours, who were referred for a second opinion from other units.

7.6 Colour Doppler Imaging

Following transvaginal B-mode assessment, the entire tumour was surveyed by Colour Doppler imaging. In the colour, spectral and power modes, the Doppler ultrasound had a frequency of 5 MHz. All colour Doppler examinations began with the same settings of the ultrasound system: the highest sensitivity for detection of colour Doppler signals was used allowing detection of blood flow velocities ≥3 cm/s, and the colour gain was set just below the background noise level to increase, as far as possible, the Doppler sensitivity for low-velocity flow detection.

A subjective semiquantitative assessment of the amount of blood flow within the examined tumour was made (colour score): a score of 1 was given when no colour could be found in the lesion, a score of 2 was given when only minimal colour could be detected, a score of 3 was given when a moderate amount of colour was present, and a score of 4 was given when the tumour appeared highly vascular, showing a large area of colour signals (Timmerman et al., 2000).
7.7 Serum CA 125 measurements

A blood sample was taken from a proportion of women (only those in whom the RMI needed to be measured or in the randomised trial) for measurement of CA-125 levels using Immuno-1 analyser (Bayer Diagnostics, Tarrytown, NY, USA). The detection limit of the assay was 1 U/mL, the coefficients of intra-assay and inter-assay variance at the levels of 32.5 U/mL and 279 U/mL were 2%, 1.1% and 4.5%, 1.1% respectively.

7.8 Histopathology and staging

A histopathologist examined specimens that were obtained at the time of surgery. Specimens were mounted onto blocks and one section per centimetre was taken for examination. Tumours were classified according to the World Health Organisation (WHO) guidelines (Serov et al., 1973). Ovarian malignancies were staged according to the classification of the International Federation of Gynaecology and Obstetrics (FIGO) (Shepherd, 1989).

7.9 Statistical Analysis

Database files were set up using Microsoft Excel (Redmond, Washington, USA) for windows to facilitate data entry and retrieval. Statistical analyses were carried out using Medcalc® version 9.2.0.2 and version 9.2.0.3 (Medcalc Software, Mariakerke, Belgium) or SPSS version 11 (SPSS Inc., Chicago, Illinois, USA) or the SAS system 9.1.3 (SAS Institute Inc., Cary, NC, USA). Continuous variables were expressed as mean and SD or median and interquartile range. Dichotomous variables were described using proportions. When suitable, relative effect sizes were quoted with difference between medians or relative risk (RR) with 95% confidence intervals. The statistical significance of differences in continuous variables was determined using Mann-Whitney U-test,
McNemar’s test, Kruskall-Wallis test or Student t-test depending on data distribution. Proportions were compared using Yates' corrected Chi-square test or Fisher’s exact test. Two-tailed p<0.05 was considered statistically significant. The diagnostic accuracy of the tests was assessed using sensitivity, specificity, positive and negative predictive value (PPV, NPV) and positive (LR+) and negative (LR-) likelihood ratio measures. The likelihood ratio of a positive test (sensitivity/[1-specificity]) and the likelihood ratio of a negative test ([1-sensitivity]/specificity) were also calculated in order to eliminate the effect of the prevalence of ovarian malignancy on the interpretation of the value of the diagnostic test (Chien and Khan, 2001). LR+ >10 or LR- <0.1 indicate a very useful test, 5-10 and 0.1-0.2 indicate a moderately useful test, 2-5 and 0.2-0.5 indicate a somewhat useful test, and 1-2 and 0.5-1 indicate a test of little usefulness (Chien and Khan, 2001; Jaeschke R et al., 1994). Diagnostic performance of the tests was also assessed by examining the areas under receiver-operating characteristics (ROC) curves.

Percentage agreement was calculated and the agreement beyond chance was estimated using weighted Cohen’s kappa. Kappa values of 0.81–1.0 were assumed to indicate very good agreement, kappa values of 0.61–0.80 good agreement, kappa values of 0.41–0.60 moderate agreement, Kappa values of 0.21–0.40 fair agreement and Kappa values < 0.20 poor agreement (Brennan and Silman, 1992).
7.10 Ethics Committee approval

The research was approved by the ethics committee and the research and development committee at King’s College Hospital and at Guy’s and St Thomas’ NHS Foundation Trust and it was entered in the registry of randomised trials. All patients gave informed consent to participate in the study. A proportion of patients were recruited through the on-going-familial ovarian cancer screening programme, which has received prior ethical approval. Another proportion of patients included formed a part of the multicentre International Ovarian Tumor Analysis (IOTA) group collaboration, which was approved by the King’s College Hospital Ethics Committee.
Chapter 8

Study 1

A comparative study of the risk of malignancy index and the ovarian crescent sign for the diagnosis of invasive ovarian cancer

8.1 Background

In recent years, ultrasound has been increasingly used for the assessment of women presenting with a wide range of gynaecological complaints. During a pelvic scan, an attempt is routinely made to identify the ovaries and to examine their morphological appearance. This has led to a substantial increase in the number of asymptomatic women diagnosed with ovarian abnormalities (Borgfeldt and Andolf, 1999). Although, the quality of ultrasound equipment has significantly improved in recent years, there is still no consensus on the most effective diagnostic criteria to discriminate between benign and malignant ovarian tumours on ultrasound scan (Granberg et al., 1990; Sassone et al., 1991; Lerner et al., 1994; Jacobs et al., 1990; Tailor et al., 1997; Timmerman et al., 1999a; Valentin et al., 2001; Aslam et al., 2000a). As a result women are often referred for additional investigations, such as tumour markers and magnetic resonance imaging (MRI), in order to clarify the nature of ovarian lesions. Many women are also offered surgery because of concerns about the possibility of malignancy, which is costly and carries the risk of complications.

The risk of malignancy index (RMI) - a scoring system based on a combination of demographic and ultrasound data with measurement of serum CA-125 - has been widely adopted in the UK to facilitate triage of women with ovarian tumours for referral to tertiary gynaecological oncology units (Jacobs et al., 1990; RCOG greentop guidelines, 2003). Although the RMI is relatively a simple test to use in clinical practice, its false negative and false positive rates are significant (Aslam et al., 2000a).
In addition, the results are not immediately available and the test is relatively costly, as it involves the use of serum biochemistry.

Recent studies have shown that a detailed sonographic examination of an ovarian tumour by an expert in gynaecological ultrasonography may be superior to all previously described diagnostic models for noninvasive diagnosis of ovarian cancer (Timmerman et al., 1999b). However, in routine clinical practice most scans are performed by sonographers, many of whom have no particular expertise in the characterisation of ovarian tumours.

We have recently described a new morphological ultrasound feature, the ‘ovarian crescent sign’ (OCS), which depends on the fact that healthy ovarian tissue can be seen adjacent to the cyst within the ipsilateral ovary (Hillaby et al., 2004). The initial report showed that this morphological ultrasound sign has a potential of becoming a simple and effective way of excluding an invasive ovarian malignancy, without the need for a detailed morphological assessment of the tumour or the use of serum biochemistry.

In this study, we prospectively compared the value of the OCS and that of RMI for the diagnosis of ovarian cancer in a large group of women with an ultrasound diagnosis of ovarian tumour.

8.2 Methods
This prospective observational study was carried out at the Gynaecological Assessment Unit at King’s College Hospital, which is a tertiary referral gynaecological ultrasound centre. All women referred to the unit underwent a pelvic ultrasound scan, which involved a detailed examination of the uterus and the adnexa. All examinations were performed by gynaecologists with special interest in gynaecological ultrasound using an Aloka SSD-5000 machine (Aloka Co, Tokyo, Japan). Transvaginal and/or transabdominal scans were performed as described in section 7.5. All women with a
diagnosis of an ovarian tumour were invited to join the study. Simple cysts, defined as anechoic unilocular lesions with regular walls and no internal septations or solid parts (Timmerman et al., 2000), were all excluded from the study.

Demographic data and menopausal status were recorded as described in section 7.3. The following morphological ultrasound information was recorded in each case: site and volume of cyst, presence of septations and solid areas within the cyst, metastases and ascites. Healthy ovarian tissue adjacent to the tumour – the OCS- was defined as visible hypoechoogenic tissue with or without ovarian follicles enclosed within the ovarian capsule and located adjacent to the cyst wall. This tissue would not be separated from the cyst when applying a moderate amount of pressure (Figure 2) (Hillaby et al., 2004).

A blood sample was taken from each patient to assess the serum CA125 levels with the Immuno-1 analyser (Bayer Diagnostics, Tarrytown, NY, USA). The RMI was calculated as originally described by Jacobs et al. (1990) (please refer to RMI calculation in chapter 3, paragraph 3.4.1).

In women with ovarian tumours, indications for surgery were suspected ovarian malignancy, clinical symptoms attributable to the adnexal cyst or patient’s request to have the ovarian tumour removed. Women who did not have an indication for surgical treatment were managed expectantly and were invited to attend for a follow-up ultrasound scan 6-8 weeks later. The cyst was classified as benign on follow up visit if the woman remained asymptomatic and there was no increase in the size of the cyst and these women attended for another follow-up scan 3 months later. All women in whom the cyst increased in size were offered surgery. However, cysts that resolved spontaneously were classified as being functional in nature. In women who underwent surgery the ultrasound findings were compared to the histology, which was classified according to the World Health Organisation (WHO) guidelines (Serov et al., 1973).
Ovarian malignancies were staged according to the classification of the International Federation of Gynaecology and Obstetrics (FIGO) (Shepherd, 1989).

The study was approved by the King's Research Ethics Committee and all patients gave their informed consent.

8.3 Statistical Analysis
All statistical analyses were carried out using SPSS version 11 (SPSS Inc., Chicago, Illinois, USA). The statistical significance of differences in continuous variables was determined using Mann-Whitney, Kruskall-Wallis or Student’s t-tests depending on the distribution of the data. Proportions were compared using Yates corrected $\chi^2$ test. Two-tailed $P < 0.05$ was considered statistically significant.

The diagnostic accuracy of the tests was assessed using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and positive and negative likelihood ratio (LR+/LR-) measures. Diagnostic performance of the tests was also assessed by examining the areas under receiver-operating characteristics (ROC) curves.

A minimum of 90 patients was necessary to detect a 22% difference between the previously reported sensitivities, in diagnosing invasive ovarian malignancy, of 74% for RMI (Aslam et al., 2000a) and 96% for ovarian crescent sign (Hillaby et al., 2004), with 85% power.

8.4 Results
One hundred and six consecutive women with ovarian masses were included in the study. The indications for the ultrasound scan were: pain in 38 (35.9%) women, suspected ovarian tumour on a previous routine scan 37 (34.9%) women, menstrual disorders in 17 (16%) women, clinical suspicion of abdominal mass in seven (6.6%)
women, ovarian cancer screening in four (3.8%) women and postmenopausal bleeding in three (2.8%) women. 55 (51.9%) women were managed expectantly and 51 (48.1%) women had surgery. Thirty-one (60.8%) of the women operated on had minimally invasive surgery, while 20 (39.2%) had a laparotomy. The indications for surgery in all 51 women operated on were: symptoms attributable to the adnexal lesion in 25 (49.0%) women, suspected malignancy in 14 (27.5%) women and patient’s choice in 12 (23.5%) women. The final diagnoses are listed in Table 2. There were significant differences in types of tumours in women who were managed expectantly in comparison to those who underwent surgery ($\chi^2 = 89.86; p < 0.001$)

None of the women had invasive tumour on one side and non-invasive one on the other. FIGO staging and histological types of invasive cancers are listed in Table 3.

A RMI > 200 was found in 8/9 (88.9%) women with invasive ovarian cancer and in 8/92 (8.7%) of those with benign lesions. RMI was < 200 in one case of invasive cancer, which was a stage II clear cell carcinoma. None of the five women with borderline tumours had RMI above that threshold. The ovarian crescent sign was absent in all nine cases of invasive ovarian cancer, 3/5 (60%) of borderline tumours and 4/92 (4.3%) of benign cysts.

False positive results with the RMI typically occurred in women with endometriotic cysts, whilst the OCS was more likely to provide a false positive result in women with benign cystadenomas (Table 4).

A RMI > 200 diagnosed invasive ovarian malignancy with a sensitivity of 89% (95% CI, 0.57 – 0.98), specificity of 92% (95% CI, 0.85 – 0.96), PPV of 50%, NPV of 99%, LR+ of 10.78 (95% CI, 5.34 – 21.77) and LR- of 0.12 (95% CI, 0.02 – 0.77). Absence of the OCS gave a sensitivity of 100% (95% CI, 0.70 – 1.0), specificity of 93% (95% CI, 0.86 – 0.96), PPV of 56%, NPV of 100%, LR+ of 13.86 (95% CI, 6.79 –
28.29) and LR- of 0, which was significantly better than the RMI (p<0.01). The area under the ROC curve for the OCS was 0.928 (SE 0.052, asymptotic 95% CI 0.826 - 1.030) vs. 0.822 (SE 0.081, asymptotic 95% CI 0.62 - 0.981) for RMI > 200 (Figure 3).
Table 2. Types of ovarian lesions and management strategies applied in each group.

<table>
<thead>
<tr>
<th>Final diagnosis</th>
<th>Laparoscopy N (%)*</th>
<th>Laparotomy N (%)*</th>
<th>Expectant N (%)†</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dermoid</td>
<td>16 (45.7)</td>
<td>3 (8.6)</td>
<td>16 (45.7)</td>
<td>35</td>
</tr>
<tr>
<td>Cystadenoma</td>
<td>5 (62.5)</td>
<td>3 (37.5)</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>Endometrioma</td>
<td>3 (13.6)</td>
<td>1 (4.6)</td>
<td>18 (81.8)</td>
<td>22</td>
</tr>
<tr>
<td>Cystadenofibroma</td>
<td>2 (100)</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Fibroma</td>
<td>1 (50)</td>
<td>1 (50)</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Functional cyst</td>
<td>2 (8.7)</td>
<td>-</td>
<td>21 (91.3)</td>
<td>23</td>
</tr>
<tr>
<td>Borderline</td>
<td>2 (40)</td>
<td>3 (60)</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>Epithelial cancer</td>
<td>-</td>
<td>8 (100)</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>Non-epithelial cancer</td>
<td>-</td>
<td>1 (100)</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
<td>20</td>
<td>55</td>
<td>106</td>
</tr>
</tbody>
</table>

* Final diagnosis based on histological report. † Final diagnosis based on ultrasound appearance.
Table 3. Histology and International Federation of Gynaecology and Obstetrics stage of invasive ovarian cancer.

<table>
<thead>
<tr>
<th></th>
<th>Stage I (%)</th>
<th>Stage II (%)</th>
<th>Stage III (%)</th>
<th>Stage IV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epithelial</td>
<td>1 (12.5)</td>
<td>3 (37.5)</td>
<td>3 (37.5)</td>
<td>1 (12.5)</td>
</tr>
<tr>
<td>Non-epithelial</td>
<td>0</td>
<td>1 (100)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 4. False-positive cases in the diagnosis of invasive ovarian cancer by the risk of malignancy index (RMI) and ovarian crescent sign (OCS).

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>RMI &gt; 200 (n)</th>
<th>Negative OCS (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional ovarian cyst</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Dermoid</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Cystadenoma</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Endometrioma</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Borderline tumour</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
Figure 2 A complex ovarian cyst with healthy ovarian tissue (arrows) clearly visible adjacent to the wall of the cyst (the ovarian crescent sign).
Figure 3 Receiver-operating characteristics curve comparing the diagnostic performance of the risk of malignancy index (RMI) and the ovarian crescent sign (OCS) for the diagnosis of invasive ovarian cancer. Solid line, RMI >200; Dashed line, OCS.
Chapter 9

Study 2
The accuracy of ultrasound subjective “pattern recognition” for the diagnosis of borderline ovarian tumours

9.1 Background
An accurate pre-operative differential diagnosis of ovarian tumours is of major clinical significance due to the implications for further management and survival. In the past, ovarian tumours used to be classified only as benign or malignant. Since 1971, the International Federation of Gynecology and Obstetrics (FIGO) (1971) has classified borderline ovarian tumours as an entity separate from benign and invasive ovarian tumours (Serov et al., 1973). However, until recently, borderline ovarian tumours (BOT) have been treated in the same way as invasive malignant tumours. BOTs are more prevalent in women of childbearing age (Gotlieb et al., 2005, Gotlieb et al., 1998; Boran et al., 2005) and their prognosis is generally good (Ahmed and Lawton, 2005; Silverberg et al., 2004). They often follow a relatively benign course (Ahmed and Lawton, 2005; Silverberg et al., 2004) and therefore fertility sparing conservative surgery is often contemplated in women with BOT who wish to preserve their reproductive potential (Maneo et al., 2004). BOTs are divided into two major histological groups: serous (two subtypes: typical serous and micropapillary) and mucinous (two subtypes: gastrointestinal and endocervical) (Ahmed and Lawton, 2005).

In the past decade, ultrasound examination has been accepted as the optimal diagnostic modality for the non-invasive assessment of ovarian tumours. However, the differentiation between various types of ovarian tumours on ultrasonography is sometimes difficult. Although, several ultrasound-based diagnostic algorithms have been proposed to facilitate differentiation between benign and malignant tumours
(Granberg et al., 1990; Sassone et al., 1991; Lerner et al., 1994; Jacobs et al., 1990; Tailor et al., 1997; Timmerman et al., 1999a; Aslam et al., 2000a), the "pattern recognition" method remains the best way of assessing the nature of ovarian tumours (Valentin et al., 2001). However, the accuracy of pattern recognition is dependent on the operator's skill and experience (Timmerman et al., 1999b). Several recent studies examined the gray-scale sonographic and colour Doppler features of borderline ovarian tumours (Exacoustos et al., 2005; Pascual et al., 2002; Wu et al., 1994). There were also attempts to describe the typical sonographic features of different histopathological subtypes of BOTs (Fruscella et al., 2005; Exacoustos et al., 2005; Gotlieb et al., 2000). However, the diagnostic accuracy of pattern recognition for the differential diagnosis between benign, borderline and invasive ovarian tumours has not been tested as yet.

The aim of this study was to prospectively assess the value of pattern recognition for the differential diagnosis of adnexal tumours and in particular its accuracy in establishing the specific diagnosis of histological subtypes of BOTs.

9.2 Methods
This was a prospective observational study of women who were referred to our regional gynaecological cancer centre with the diagnosis of an adnexal mass. They were all assessed clinically and had a level II (EFSUMB Newsletter, 2005) gynaecological ultrasound scan. Women with conclusive features of ovarian cancer or benign tumours were managed according to the centre's protocols. Those women, in whom the nature of the lesion was uncertain during the level II gynaecological ultrasound scan, were referred for a level III (EFSUMB Newsletter, 2005) ultrasound scan at our centre, which is a tertiary referral gynaecological ultrasound unit. All women referred to the unit underwent a pelvic ultrasound scan, which involved a detailed examination of the uterus and the adnexa. All examinations were performed by gynaecologists with special
interest in gynaecological ultrasound using an Aloka SSD-5000 machine (Aloka Co, Tokyo, Japan). Transvaginal and / or transabdominal scans were performed as described in section 7.5.

Demographic data and menopausal status were recorded as described in section 7.3. The following morphological ultrasound information was recorded in each case: site and volume of cyst, presence of septations and solid areas within the cyst, metastases and ascites. The tumour pattern recognition method was used prospectively to differentiate between various types of ovarian tumours.

Morphological features suggestive of a BOT were: unilocular cyst with a positive ovarian crescent sign and extensive papillarities arising from the inner wall, or a cyst with a well-defined multilocular nodule. Solid papillary projections were defined as any solid projections into the cyst cavity arising from the cyst wall with a height greater than or equal to 3 mm. A multilocular nodule (“honeycomb nodule”) was defined as a predominantly solid nodule with cystic areas arising from the inner cyst wall. The presence of multiple papillary projections and positive ovarian crescent sign were suggestive of a serous type (Figure 4) or an endocervical type BOT (Figure 5), while the presence of a “Honeycomb nodule” and thick echogenic fluid content was indicative of a gastrointestinal (GI) type BOT (Figure 6).

The presence of healthy ovarian tissue adjacent to the tumour “the ovarian crescent sign” (Hillaby et al. 2004) was used to exclude invasive ovarian cancer. The ovarian crescent sign was defined in section 8.2.

Women who were diagnosed with benign ovarian tumours underwent conservative surgery. Women of childbearing age with a suspected BOT, who wished to preserve their fertility, underwent fertility sparing surgery such as cystectomy, oophorectomy or adnexectomy. Other women with BOTs who completed their families and those with suspected invasive ovarian cancer underwent a laparotomy, total
abdominal hysterectomy, bilateral salpingo-oophorectomy and omentectomy at the gynaecological cancer centre. The ultrasound findings were compared to the histology, which was classified according to the World Health Organisation (WHO) guidelines (Serov et al., 1973). Ovarian malignancies were staged according to the classification of the International Federation of Gynaecology and Obstetrics (FIGO (International Federation of Gynaecology and Obstetrics, 1971; Shepherd, 1989).

The study was approved by the King’s Research Ethics Committee and all patients gave informed consent.

9.3 Statistical Analysis
All statistical analyses were carried out using SPSS version 14 (SPSS Inc., Chicago, Illinois, USA). The statistical significance of differences in continuous variables was determined using Mann-Whitney or Student t-tests depending on data distribution. Proportions were compared using Yates corrected $\chi^2$ test. Two-tailed p value of <0.05 was considered statistically significant. The diagnostic accuracy of the tests was assessed using sensitivity, specificity and likelihood ratio (LR) measures.

9.4 Results
A total of 224 women with an adnexal mass of uncertain nature were referred for an expert scan, 166 (74.1%) of whom underwent surgery. The remaining 58 (25.9%) had functional ovarian cysts, hydrosalpinges, peritoneal inclusion cysts or uterine fibroids diagnosed on ultrasound scan, and these did not require surgical intervention. Women with suspected functional ovarian cyst had at least one follow up ultrasound scan to confirm spontaneous resolution of the cyst.

Histopathological results were available in all 166 women who underwent surgery and they were all included in the final data analyses. There were 99 (59.6%,
95% CI 52 – 67) benign tumours, 35 (21.1%, 95% CI 15 – 27) borderline and 32 
(19.3%, 95% CI 13 – 25) invasive ovarian cancers. The patients’ mean age was 42 years 
(range, 14 – 88) and 45/166 (27%, 95% CI 20 – 34) women were postmenopausal. 
Patients’ demographic data and clinical symptoms at presentation are listed in Table 5. 
When comparing women’s demographic data and clinical symptoms, there were 
statistically significant differences between benign and invasive ovarian tumours and 
also between borderline and invasive ovarian tumours (except for the menopausal 
status), but not between benign and borderline ovarian tumours.

FIGO staging and histological types of BOT and invasive cancers are listed in 
Table 6; 25/35 (71.4%, 95% CI 56 – 86) BOT underwent surgical staging. All the 
staged GI type BOTs and the majority of the staged serous / endocervical type BOTs 
were FIGO stage I. Only 2/16 (12.5%, 95% CI 0 – 29) serous / endocervical type BOTs 
were FIGO stage III.

Morphological ultrasonic characteristics and final histology of all ovarian 
tumours in the study population are summarised in Table 7. Papillary projections were 
present in approximately half of all borderline and invasive ovarian tumours, and in one 
fifth of ovarian cystadenomas ($\chi^2 = 35.0$, p < 0.001). Solid papillary projections were 
much more frequent in serous / endocervical type BOTs (80%, 95% CI 63 - 98) than in 
cases of GI type BOTs (7%, 95% CI 0 – 20) ($\chi^2 = 18.5$, p < 0.001). There were also 
significant differences in the presence of the ovarian crescent sign between benign, 
borderline and invasive ovarian tumours ($\chi^2 = 75.6$, p < 0.001). In the subtypes of BOTs, 
the ovarian crescent sign was present in 75% (95% CI, 56 – 94) serous / endocervical 
type BOTs and in 20% (95% CI, 0 – 40) GI type BOTs ($\chi^2 = 10.38$, p < 0.01). Thick 
fluid was present in all GI type BOT. However, this was not a specific finding as it was 
also found in most endometriomas as well as in some serous / endocervical type BOT 
and occasionally in epithelial ovarian cancers. The honeycomb nodule was a highly
specific feature of GI type BOTs. However, its sensitivity was low as it was absent in nearly half of cases.

The false-negative and the false-positive cases encountered in this study are summarised with their respective gray scale morphological appearances in Table 8. The women's age and the tumour volume were not significantly different in cases of true-positive, false-positive or false-negative diagnoses in women with different subtypes of BOTs (Table 9). The median volume of the false-negative unilocular cysts was 1226 (range 412 – 5661) mL.

The area under the receiver-operating characteristics curve, sensitivity, specificity, and positive and negative likelihood ratios of pattern recognition combining different morphological features for the diagnosis of benign, borderline or malignant tumours are shown in Table 10. The sensitivity was better in cases of serous / endocervical type (75%, 95% CI 56 – 94) compared to GI type (60%, 95% CI 35 – 85) BOTs.
Table 5. Demographic data and clinical symptoms at presentation of 166 women with an adnexal mass.

<table>
<thead>
<tr>
<th>Type of adnexal mass</th>
<th>Benign n=99</th>
<th>Borderline n=35</th>
<th>Invasive N=32</th>
<th>p value (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years, mean (range))</td>
<td>39 (14-83)</td>
<td>39 (21-65)</td>
<td>52 (17-88)</td>
<td>0.84</td>
</tr>
<tr>
<td>Postmenopausal (n %) [95% CI]</td>
<td>19 (19) [11-27]</td>
<td>7 (20) [7-33]</td>
<td>19 (59) [42-76]</td>
<td>0.89</td>
</tr>
<tr>
<td>Pain (n %) [95% CI]</td>
<td>41 (41) [31-51]</td>
<td>9 (26) [12-41]</td>
<td>5 (16) [3-29]</td>
<td>0.15</td>
</tr>
<tr>
<td>Abdominal distension (n %) [95% CI]</td>
<td>8 (8) [3-13]</td>
<td>5 (14) [3-26]</td>
<td>18 (56) [39-73]</td>
<td>0.46</td>
</tr>
<tr>
<td>Incidental finding on USS or clinical examination (n %) [95% CI]</td>
<td>50 (51) [41-61]</td>
<td>21 (60) [44-76]</td>
<td>9 (28) [12-44]</td>
<td>0.44</td>
</tr>
</tbody>
</table>
Table 6. Histology and International Federation of Gynaecology and Obstetrics stages of borderline and invasive ovarian cancers.

<table>
<thead>
<tr>
<th></th>
<th>Stage I</th>
<th>Stage II</th>
<th>Stage III</th>
<th>Stage IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>Serous / endocervical</td>
<td>14 (87.5)</td>
<td>0 (0)</td>
<td>2 (12.5)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>BOT* (n= 16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI BOT* (n=9)</td>
<td>9 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Epithelial invasive</td>
<td>9 (37.5)</td>
<td>6 (25)</td>
<td>6 (25)</td>
<td>3 (12.5)</td>
</tr>
<tr>
<td>(n=24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-epithelial invasive</td>
<td>3 (37.5)</td>
<td>1 (12.5)</td>
<td>3 (37.5)</td>
<td>1 (12.5)</td>
</tr>
<tr>
<td>(n=8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*25/35 (71%) borderline tumours had formal surgical staging. BOT, borderline ovarian tumour; GI, gastrointestinal.
Table 7. Gray-scale ultrasound characteristics and histology of the tumours of the study population.

<table>
<thead>
<tr>
<th></th>
<th>Benign (n=99)</th>
<th>Borderline (n=35)</th>
<th>Invasive (n=32)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dermoid</td>
<td>Cystadenoma</td>
<td>Endometrioma</td>
</tr>
<tr>
<td>Patients (n)</td>
<td>35</td>
<td>36</td>
<td>17</td>
</tr>
<tr>
<td>Unilocular (n (%))</td>
<td>11 (31.4)</td>
<td>14 (38.9)</td>
<td>12 (70.1)</td>
</tr>
<tr>
<td>Unilocular solid (n (%))</td>
<td>15 (42.9)</td>
<td>0 (0)</td>
<td>1 (5.9)</td>
</tr>
<tr>
<td>Multilocular (n (%))</td>
<td>3 (8.6)</td>
<td>20 (55.6)</td>
<td>4 (23.5)</td>
</tr>
<tr>
<td>Multilocular solid (n (%))</td>
<td>5 (14.3)</td>
<td>2 (5.6)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Solid (n (%))</td>
<td>1 (2.9)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Papillary Projections (n (%))</td>
<td>0 (0)</td>
<td>7 (19.4)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Thick fluid content (%)</td>
<td>0 (0)</td>
<td>3 (8.3)</td>
<td>16 (94.1)</td>
</tr>
<tr>
<td>Honeycomb nodule (%)</td>
<td>0 (0)</td>
<td>1 (2.8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Ovarian crescent (%)</td>
<td>33 (94.3)</td>
<td>28 (77.8)</td>
<td>16 (94.1)</td>
</tr>
</tbody>
</table>

BOT, borderline ovarian tumour; GI, gastrointestinal.
Table 8. Ultrasound morphological appearance of false positive and false negative cases.

<table>
<thead>
<tr>
<th>Histology</th>
<th>Ultrasound diagnosis</th>
<th>Ultrasound morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 7</td>
<td>Cystadenoma</td>
<td>Serous / endocervical BOT Unilocular solid cyst with numerous PP</td>
</tr>
<tr>
<td>n = 1</td>
<td>Cystadenoma</td>
<td>GI BOT Unilocular cyst with a Honeycomb nodule</td>
</tr>
<tr>
<td>n = 3</td>
<td>Serous / endocervical</td>
<td>Cystadenoma Unilocular cyst</td>
</tr>
<tr>
<td>n = 1</td>
<td>Serous / endocervical</td>
<td>Cystadenoma Unilocular cyst with a PP</td>
</tr>
<tr>
<td>n = 1</td>
<td>Serous / endocervical</td>
<td>Cystadenoma Multilocular cyst</td>
</tr>
<tr>
<td>n = 3</td>
<td>GI BOT</td>
<td>Cystadenoma Multilocular cyst</td>
</tr>
<tr>
<td>n = 1</td>
<td>GI BOT</td>
<td>Endometrioma Unilocular cyst containing fluid with ‘ground glass’ appearance</td>
</tr>
<tr>
<td>n = 1</td>
<td>GI BOT</td>
<td>Dermoid Multilocular solid cyst</td>
</tr>
<tr>
<td>n = 1</td>
<td>GI BOT</td>
<td>Invasive cancer Multilocular solid cyst</td>
</tr>
</tbody>
</table>

BOT, borderline ovarian tumour; GI, gastrointestinal; PP, papillary projection.
Table 9. Age of women and tumour volume in true positive, false positive and false negative cases of borderline ovarian tumour (BOT).

<table>
<thead>
<tr>
<th></th>
<th>True Positive</th>
<th>False Positive</th>
<th>False Negative</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GI type BOT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>8</td>
<td>1</td>
<td>6</td>
<td>NS</td>
</tr>
<tr>
<td>Age (years, median)</td>
<td>39</td>
<td>62</td>
<td>41</td>
<td>NS</td>
</tr>
<tr>
<td>(range)</td>
<td>(22–64)</td>
<td>(31–63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour volume (mL, median (range))</td>
<td>1124</td>
<td>2469</td>
<td>476</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>(603–6401)</td>
<td>(90–5611)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Serous / endocervical types BOT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>16</td>
<td>7</td>
<td>5</td>
<td>NS</td>
</tr>
<tr>
<td>Age (years, median)</td>
<td>36</td>
<td>62</td>
<td>45</td>
<td>NS</td>
</tr>
<tr>
<td>(range)</td>
<td>(21–47)</td>
<td>(16–80)</td>
<td>(31–65)</td>
<td></td>
</tr>
<tr>
<td>Tumour volume (mL, median (range))</td>
<td>77</td>
<td>613</td>
<td>321</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>(1–2579)</td>
<td>(25–657)</td>
<td>(45–1593)</td>
<td></td>
</tr>
</tbody>
</table>

GI, gastrointestinal; NS, not significant.
Table 10. Accuracy of pattern recognition for the diagnosis of different types of ovarian tumours.

<table>
<thead>
<tr>
<th>Type of tumour</th>
<th>Area under ROC curve (95% CI)</th>
<th>Standard error</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR- (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td>0.872</td>
<td>0.031</td>
<td>0.91 (0.83–0.95)</td>
<td>0.86 (0.75–0.92)</td>
<td>6.45 (3.71–11.19)</td>
<td>0.11 (0.05–0.21)</td>
</tr>
<tr>
<td>Borderline</td>
<td>0.812</td>
<td>0.049</td>
<td>0.69 (0.52–0.81)</td>
<td>0.94 (0.88–0.97)</td>
<td>11.3 (5.53–22.8)</td>
<td>0.34 (0.21–0.55)</td>
</tr>
<tr>
<td>Malignant</td>
<td>0.977</td>
<td>0.019</td>
<td>0.97 (0.84–0.99)</td>
<td>0.99 (0.96–0.99)</td>
<td>129.8 (18.4–915.7)</td>
<td>0.03 (0.01–0.22)</td>
</tr>
</tbody>
</table>

Sensitivities, specificities and positive and negative likelihood ratios (LR+ and LR-) were calculated separately for each of the three groups of tumours.

ROC, receiver–operating characteristics.
Figure 4 Ultrasound image showing a serous borderline ovarian tumour with extensive papillary projections arising from the cyst wall. Note the presence of healthy ovarian tissue, the ‘ovarian crescent sign’ (arrow).
Figure 5 Ultrasound image showing a mucinous endocervical-type borderline ovarian tumour (BOT). Note the resemblance to a serous BOT, but more organized papillary projections. The ovarian crescent sign is also a common feature (arrow).
Figure 6 Ultrasound image showing a mucinous gastrointestinal-type borderline ovarian tumour. A ‘honeycomb nodule’ is seen suspended within the cyst cavity (arrow).
Chapter 10

Study 3

Real-time ultrasound versus evaluation of static images in the preoperative assessment of adnexal masses

10.1 Background

Trying to make an accurate prediction of the benign or malignant nature of an adnexal mass during a preoperative ultrasound examination remains an important task because the preoperative diagnosis influences the treatment strategy and, as a consequence, the prognosis of the patient (Buchweitz et al., 2005; Medeiros et al., 2005; Carley et al., 2002; Vergote et al., 2001; Hacker et al., 1983). As mentioned in previous chapters, several reports have shown that subjective evaluation of ultrasound findings (pattern recognition) by an expert sonologist is one of the best methods for discriminating between benign and malignant adnexal masses before surgery (Valentin, 1999; Timmerman et al., 1999b; Timmerman, 2004; Valentin, 2004)). If the evaluation is performed by a less experienced sonologist, the diagnostic performance is reduced (Timmerman et al., 1999b). In training hospitals, a patient with an adnexal mass is often first scanned by a junior or senior ultrasound operator. Subsequently, the images are discussed with a supervising ultrasound operator, who will often rescan the patient, because he or she believes that this will help in gaining the right information to make a correct diagnosis. However, to the best of our knowledge, it has never been confirmed in a scientific study that real-time scanning is superior to evaluating saved static images of an adnexal mass.

The aim of this study is to determine if the prediction of the nature of an adnexal mass using pattern recognition, i.e. subjective evaluation of gray-scale and Doppler
ultrasound findings, is as accurate when based on static images as it is when based on a real-time ultrasound examination.

10.2 Methods

For this prospective observational study, a dataset of non-consecutive patients with an adnexal mass was created in the Early Pregnancy and Gynaecology unit of King’s College Hospital, London, UK. Patients were included if they had complex or difficult-to-classify masses when using pattern recognition. A large number of patients were referred to this tertiary referral centre after they had undergone an ultrasound examination in their referring hospital. In this way we collected a dataset of 171 selected patients with adnexal masses. The study was a part of the multicentre IOTA (International Ovarian Tumor Analysis) group collaboration, which was approved by the King’s College Hospital Ethics Committee.

All patients included in this study were preoperatively scanned by one of the expert sonologists at the tertiary referral centre, the “real-time” sonologist (DJ). The ultrasound examination was performed transvaginally and transabdominally using the gray-scale and colour/power Doppler modes of an Aloka SSD-5000 ultrasound machine (Aloka Co., Tokyo, Japan), and the images were saved electronically. The anonymised electronic images of all 171 adnexal masses were later independently evaluated by three expert sonologists, the “image experts” (DT, ACT, LV). Patients were excluded if one of the “image experts” found the quality of the images insufficient to make a reliable diagnosis. The image experts received relevant clinical information and information on the colour score if the colour content of the tumour scan was not clearly demonstrated in the static ultrasound images. The colour score is described in section 7.6. Both the real-time sonologist and the image experts classified each mass as benign or malignant using pattern recognition. Borderline ovarian tumours were regarded as malignant. They also
stated with which degree of confidence they made their diagnosis (certainly benign or malignant, probably benign or malignant, uncertain). All four ultrasound experts involved in this study had more than 10 years of experience in gynaecologic ultrasound at the start of the study, and they are all senior clinicians in tertiary referral centres.

The primary outcome was the histopathological classification of the adnexal mass following the WHO guidelines (Serov et al., 1973). In case of malignancy, patients were staged according to the FIGO criteria (Shepherds, 1989).

10.3 Statistical Analysis
All statistical analyses were carried out using SAS version 9.1.3 for Windows (SAS Institute Inc., Cary, NC, USA). The accuracy, sensitivity and specificity were calculated for the real-time sonologist and for each of the image experts. The performance of the real-time expert was compared with the performance of each image expert and also with the “consensus opinion” of the three image experts. The consensus opinion was defined as the diagnosis predicted by at least two of the three image experts. McNemar’s test was used to determine the statistical significance of a difference in nominal variables between matched samples.

10.4 Results
Five of the 171 masses were excluded because of poor image quality, leaving 166 adnexal masses for evaluation. Seventy masses (42%) were malignant of which 34 (49%) were borderline tumours. Ninety-six masses (58%) were benign. Table 11 shows the histopathological diagnoses.

Table 12 shows the performance of the real-time sonologist, that of each image expert and that of the consensus opinion of the image experts. The accuracy and specificity of the real-time sonologist were superior to those of two of the image experts
and also to the consensus opinion of the three image experts, although the latter was only significant with respect to specificity.

The consensus opinion had nine more false positive cases in comparison with the real time sonologist reducing the specificity significantly. These false positive cases were one dermoid, two fibromas and six cystadenomas. The fibromas were presumed to be rare malignant tumours by the image experts and the cystadenomas and the dermoid were presumed to be borderline ovarian tumours. In only one of these nine cases, two of three image experts were very confident about their diagnosis of malignancy, and in three cases one image expert was certain. In the remaining five false positive cases the image experts were either completely unsure or not really confident (probably benign or probably malignant). The real-time expert stated that he was very confident of his benign diagnosis in four of the nine cases and stated that the diagnosis was probably benign in the remaining five. The four cases that were misclassified by both the consensus opinion and the real-time expert were three cystadenofibromas (one with Brenner tumour component), all of which were misclassified as borderline tumours, and one thecoma that was misclassified as a rare malignant tumour by all the experts.
Table 11. Histopathological diagnoses of the adnexal masses included in the study.

<table>
<thead>
<tr>
<th>Histopathological diagnosis</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benign</strong></td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td>96 (57.8)</td>
</tr>
<tr>
<td>Dermoid cyst</td>
<td>35 (21.1)</td>
</tr>
<tr>
<td>Cystadenoma/-fibroma</td>
<td>35 (21.1)</td>
</tr>
<tr>
<td>Endometrioma</td>
<td>16 (9.6)</td>
</tr>
<tr>
<td>Fibroma</td>
<td>6 (3.6)</td>
</tr>
<tr>
<td>Simple cyst/functional cyst</td>
<td>2 (1.2)</td>
</tr>
<tr>
<td>Abscess</td>
<td>1 (0.6)</td>
</tr>
<tr>
<td>Rare benign tumour</td>
<td>1 (0.6)</td>
</tr>
<tr>
<td><strong>Malignant</strong></td>
<td></td>
</tr>
<tr>
<td>Mucinous borderline tumour</td>
<td>16 (9.6)</td>
</tr>
<tr>
<td>Serous borderline tumour</td>
<td>18 (10.9)</td>
</tr>
<tr>
<td>Primary invasive carcinoma</td>
<td>25 (15.1)</td>
</tr>
<tr>
<td>Rare malignant tumour</td>
<td>11 (6.6)</td>
</tr>
</tbody>
</table>
Table 12. Accuracy, sensitivity, and specificity with regard to malignancy of subjective evaluation of gray-scale and Doppler ultrasound findings in an adnexal mass during scanning ("real-time" sonologist), and by three "image experts" (A, B, C) who evaluated the static images saved by the real-time sonologist, and of the "consensus opinion" of the three image experts (i.e., the diagnosis suggested by at least two of the three experts).

<table>
<thead>
<tr>
<th></th>
<th>Real-time sonologist</th>
<th>Image expert A</th>
<th>Image expert B</th>
<th>Image expert C</th>
<th>Consensus opinion of A, B, and C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (%) (n)</td>
<td>89 (148/166)</td>
<td>89 (147/166)</td>
<td>82 (136/166)</td>
<td>83 (138/166)</td>
<td>85 (141/166)</td>
</tr>
<tr>
<td></td>
<td>p = 0.8084</td>
<td></td>
<td>p = 0.0186</td>
<td>p = 0.0198</td>
<td>p = 0.0707</td>
</tr>
<tr>
<td>Sensitivity (%) (n)</td>
<td>80 (56/70)</td>
<td>86 (60/70)</td>
<td>86 (60/70)</td>
<td>80 (56/70)</td>
<td>83 (58/70)</td>
</tr>
<tr>
<td></td>
<td>p = 0.1573</td>
<td></td>
<td>p = 0.2059</td>
<td>p = 1.0000</td>
<td>p = 0.4142</td>
</tr>
<tr>
<td>Specificity (%) (n)</td>
<td>96 (92/96)</td>
<td>91 (87/96)</td>
<td>79 (76/96)</td>
<td>84 (82/96)</td>
<td>86 (83/96)</td>
</tr>
<tr>
<td></td>
<td>= 0.0956</td>
<td></td>
<td>p &lt; 0.0001</td>
<td>p = 0.0016</td>
<td>p = 0.0027</td>
</tr>
<tr>
<td>LR-</td>
<td>0.21</td>
<td>0.16</td>
<td>0.18</td>
<td>0.23</td>
<td>0.20</td>
</tr>
</tbody>
</table>

The p-values show the statistical significance of the difference between the real-time sonologist and the others.

LR+, positive likelihood ratio. LR-, negative likelihood ratio
Chapter 11

Study 4

The use of ultrasound pattern recognition by expert ultrasound operators to identify borderline ovarian tumours: a study of diagnostic performance and interobserver agreement of ultrasound diagnoses

11.1 Background

An accurate pre-operative diagnosis of the nature of adnexal tumours facilitates the choice between conservative and surgical management. In women who are booked for surgery, information obtained by ultrasonography helps to decide on the most appropriate surgical approach and assists in deciding whether referral for specialised gynaecological oncological care is required.

Traditionally, adnexal tumours were classified on ultrasound scan as being either benign or malignant. Recent reports have suggested that, using pattern recognition, it is possible to establish a correct pre-operative diagnosis of borderline ovarian tumours with a sensitivity of 69%. The low sensitivity may be explained by variations in the morphological appearance of borderline ovarian tumours. Nevertheless, the ability to identify borderline ovarian tumours pre-operatively is of great importance. Borderline ovarian tumours tend to occur in younger women (Gotlieb et al., 2005) and they are usually associated with good long term prognosis (Ahmed and Lawton, 2005; Boran et al., 2005, Silverberg et al., 2004). Therefore, pre-operative identification of this type of ovarian tumours should allow more conservative, fertility-sparing, surgery to be carried out on women who are in the reproductive age group and who have still not completed their families. The agreement between experienced ultrasound operators with regard to identifying borderline ovarian tumours as a separate entity from benign and invasive malignant ovarian tumours has not been formally tested.
The aim of this study was to evaluate the reproducibility of ultrasound ‘pattern recognition’ for the diagnosis of borderline ovarian tumours by asking experienced ultrasound operators to systematically examine a large data set containing representative images of different types of adnexal tumours.

11.2 Methods

The database of the Early Pregnancy and Gynaecology Assessment Unit at King’s College Hospital, UK was searched to identify all women who were diagnosed with adnexal tumours in the period between January 2004 and June 2006. Only women who underwent surgery and in whom a final histological diagnosis was available were included. The cases were selected arbitrarily to ensure that the dataset included a mix of representative examples of benign, borderline and invasive malignant ovarian tumours. The study was a part of the multicentre IOTA (International Ovarian Tumor Analysis) group collaboration, which was approved by the King’s College Hospital Ethics Committee.

All original real-time examinations were performed by expert ultrasound examiners using an Aloka SSD-5000 machine (Aloka Co, Tokyo, Japan). In each case, the indication for the scan as well as the clinical history was recorded in a database (PLA- Fetal Database, Viewpoint Bildverarbeitung GmbH, Wessling, Germany). Static two-dimensional B-mode images of standard views of the uterus, ovaries, and pouch of Douglas were obtained in all cases and they were digitally stored in the database. All detectable adnexal abnormalities were measured in three orthogonal planes. In each case, an effort was made to record ultrasound features, which were considered characteristic of a particular type of adnexal tumour.

The stored static two-dimensional B-mode images were independently assessed by three expert operators (EFSUMB Newsletter, 2005) (ACT, DT and LV). Each
operator was given information about the indication for the scan and relevant clinical history, but they were unaware of the pre-operative real-time ultrasound diagnosis and of the histopathological findings. The operators were first asked to determine whether the images were of sufficient quality. If the quality of the images was found to be suboptimal by any of the three operators then the case was excluded from the final data-analysis. If the images were considered satisfactory, the operators were asked to examine them in detail and to classify the tumours as benign, borderline or invasive malignant.

The outcome measures included: accuracy of each operator in diagnosing benign, borderline or invasive malignant ovarian tumours, and the interobserver agreement in classifying the tumours as benign, borderline or invasive malignant.

11.3 Statistical Analysis

The statistical analyses were carried out using the SAS system 9.1.3, SAS Institute Inc., Cary, NC, USA.

The $\chi^2$ test, Fisher's exact test and McNemar's test were used to test the statistical significance of differences in discrete data, Mann-Whitney's test was used to compare the continuous variable (patient's age). All reported p values were two-tailed and we considered as significant a nominal $p<0.05$ although multiple tests were done, considering that all the analyses that were carried out had an exploratory purpose.

Percentage agreement was calculated and the agreement beyond chance was estimated using weighted Cohen's kappa. Kappa values of 0.81–1.0 were assumed to indicate very good agreement, kappa values of 0.61–0.80 good agreement, kappa values of 0.41–0.60 moderate agreement, Kappa values of 0.21–0.40 fair agreement and Kappa values < 0.20 poor agreement (Brennan and Silman, 1992).
11.4 Results

The static two-dimensional B-mode and Doppler images of 171 women diagnosed with adnexal tumours were collected. Five of 171 (3%) cases were excluded by the expert sonologists because of suboptimal image quality.

The 166 cases included in the analysis were assessed by the three expert operators (A, B and C). The data set included 96 (58%) benign ovarian tumours, 34 (20%) borderline ovarian tumours and 36 (22%) primary invasive ovarian tumours. Median age (range) was 57 years (13-88) in patients with a primary invasive ovarian tumour compared to 36 years (21-65) and 39 years (14-83) in women with borderline and benign masses, respectively (p<0.01).

Table 13 shows the experts’ diagnosis stratified by final histological diagnosis. The rates of correctly diagnosed masses for operators A, B and C were 74% (25/34), 71% (24/34) and 59% (20/34) in borderline tumours, 91% (87/96), 79% (76/96) and 85% (82/96) in benign tumours and 86% (31/36), 92% (33/36) and 89% (32/36) in primary invasive malignant tumours, respectively (borderline vs. benign A: p=0.02, B: p=0.35, C: p<0.01; borderline vs. invasive A: p=0.24, B: p=0.03, C: p<0.01). All three operators showed tendency to misclassify borderline tumours as benign rather than primary invasive: ratio of 8:1 in operator A, 6:1 in C and 4:1 in B. As a result of this, the diagnostic performance of the operators in making the discrimination between invasive and non-invasive (benign and borderline) tumours was better compared to the traditional discrimination of malignant (borderline and invasive) versus benign tumours: all three operators had higher sensitivity, specificity, LR+ and LR- in identifying invasive masses compared to malignant masses (borderline and invasive) (Tables 14 and 15).

The proportion of agreement between any two observers with regards to making an ultrasound diagnosis of benign, borderline and primary invasive malignant ovarian...
tumour varied from 81% to 86% (Table 16). Cohen's kappa ranged from 0.76 to 0.82 indicating a good to very good agreement between each pair of operators. The interobserver agreement between any two experts was very good when they were tested for their ability to discriminate between invasive and non-invasive ovarian tumours (Kappa 0.85 to 0.88), but poorer for the discrimination between malignant versus benign (Kappa 0.70 to 0.78). All three operators were concordant in classifying the tumours as invasive or non-invasive in 154/166 (93%) compared to 134/166 (81%) in classifying the tumours as benign or malignant (p<0.01).

The proportion of operators who correctly classified the adnexal masses is shown in Table 17. A correct classification was made by all three experts in 83% of the primary invasive cancers, 76% of the benign masses, compared to only 43% of the borderline malignant tumours (p<0.01).
Table 13. The diagnosis of individual experts in benign, borderline and primary invasive ovarian tumours.

<table>
<thead>
<tr>
<th>Histology</th>
<th>Expert</th>
<th>Benign</th>
<th>Borderline</th>
<th>Primary invasive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign (N = 96)</td>
<td>A</td>
<td>87</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>76</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>82</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Borderline (N = 34)</td>
<td>A</td>
<td>8</td>
<td>25</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>8</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>12</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>Primary invasive (N = 36)</td>
<td>A</td>
<td>2</td>
<td>3</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2</td>
<td>2</td>
<td>32</td>
</tr>
</tbody>
</table>
Table 14. Diagnostic performance of pattern recognition with regard to discriminating between invasive and non-invasive (benign and borderline) tumours.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Sensitivity (% (95% CI))</th>
<th>Specificity (% (95% CI))</th>
<th>LR+ (95% CI)</th>
<th>LR- (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>31/36 (86, 71-95)</td>
<td>126/130 (97, 92-99)</td>
<td>27.99 (10.57-74.10)</td>
<td>0.14 (0.06-0.32)</td>
</tr>
<tr>
<td>B</td>
<td>33/36 (92, 78-98)</td>
<td>124/130 (95, 90-98)</td>
<td>19.86 (9.04-43.66)</td>
<td>0.09 (0.03-0.26)</td>
</tr>
<tr>
<td>C</td>
<td>32/36 (89, 74-97)</td>
<td>122/130 (94, 88-97)</td>
<td>14.44 (7.31-28.54)</td>
<td>0.12 (0.05-0.30)</td>
</tr>
</tbody>
</table>

LR+, positive likelihood ratio; LR-, negative likelihood ratio;
Table 15. Diagnostic performance of pattern recognition with regard to discrimination between malignant (borderline and invasive malignancies) and benign tumours.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Sensitivity (% , 95% CI)</th>
<th>Specificity (% , 95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR- (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>60/70 (86, 75-93)</td>
<td>87/96 (91,83-96)</td>
<td>9.14 (4.87-47.15)</td>
<td>0.16 (0.09-0.28)</td>
</tr>
<tr>
<td>B</td>
<td>60/70 (86, 75-93)</td>
<td>76/96 (79, 70-87)</td>
<td>4.11 (2.75-6.15)</td>
<td>0.18 (0.10-0.32)</td>
</tr>
<tr>
<td>C</td>
<td>56/70 (80, 69-89)</td>
<td>82/96 (85, 77-92)</td>
<td>5.49 (3.33-9.03)</td>
<td>0.23 (0.15-0.38)</td>
</tr>
</tbody>
</table>

LR+, positive likelihood ration; LR-, negative likelihood ratio
Table 16. Interobserver agreement between each two of the three experts in classifying the tumours as benign, borderline and invasive malignant (N = 166).

<table>
<thead>
<tr>
<th>Experts</th>
<th>Proportion of agreement</th>
<th>Kappa (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ultrasound Diagnosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Benign</td>
<td>Borderline</td>
</tr>
<tr>
<td>A, B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A, C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B, C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 17. Proportion of experts who correctly classified the ovarian masses.

<table>
<thead>
<tr>
<th>Proportion of Experts</th>
<th>Benign N = 96</th>
<th>Borderline N = 34</th>
<th>Primary invasive N = 36</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N  %</td>
<td>N  %</td>
<td>N  %</td>
</tr>
<tr>
<td>0/3</td>
<td>6  6</td>
<td>2  6</td>
<td>2  6</td>
</tr>
<tr>
<td>1/3</td>
<td>8  8</td>
<td>10 29</td>
<td>2  6</td>
</tr>
<tr>
<td>2/3</td>
<td>9  9</td>
<td>7  21</td>
<td>2  6</td>
</tr>
<tr>
<td>3/3</td>
<td>73 76</td>
<td>15 44</td>
<td>30 83</td>
</tr>
</tbody>
</table>
Chapter 12

Study 5

Confidence of the ultrasound operator in making a differential diagnosis of an adnexal tumour: Effect on diagnostic accuracy and interobserver agreement

12.1 Background

Pattern recognition method is a subjective technique used for the ultrasound diagnosis of adnexal tumours (Valentin, 2004). Adnexal tumours often exhibit specific morphological features on gray-scale ultrasonography (Valentin, 2004, Valentin et al., 2006a), which could be used to foretell their histological type and nature. In typical cases, the ultrasound operator may be able to make the diagnosis with a high level of confidence. However, a number of adnexal tumours exhibit complex or unusual morphological features (Valentin et al., 2006b). As a result, these tumours are more difficult to classify. This may decrease both the diagnostic accuracy and the confidence with which the diagnosis is made. Operators do not usually report the level of confidence in reaching the diagnosis. In uncertain cases, however, many operators provide a morphological description of the adnexal tumour without committing themselves to a specific histological diagnosis.

The degree of the ultrasound operators' confidence in making a diagnosis of an ovarian tumour is likely to be determined by the level of their skill and experience. However, confidence may also, to some extent, reflect the personality of the operator and it may vary even in operators with similar knowledge and experience.

Ultrasound pattern recognition technique enables accurate differentiation between benign and malignant adnexal tumours when used by expert ultrasound operators (Valentin 1999; Timmerman et al., 1999b). However, we do not know how much the level of confidence, when making diagnosis of adnexal tumour, varies
between the operators and whether the diagnostic accuracy depends on the level of confidence.

The aim of this study was to assess the degree of confidence with which expert ultrasound operators differentiate between benign, borderline and invasive malignant ovarian tumours and to assess the effect of their diagnostic confidence on diagnostic accuracy and interobserver agreement.

12.2 Methods

The same methods of section 11.2 have been used in this study. In addition, the operators were also asked to provide the degree of confidence with which they made the diagnosis, confidence being classified into certain, probable and uncertain diagnosis. The operators were certain when they were nearly 100% sure of their diagnosis, probable when they were more than 50% sure and uncertain when they made a complete guess.

The main outcome measure was the degree of confidence of the three operators in making the diagnosis of benign, borderline or invasive malignant ovarian tumour. Secondary outcomes included:

- Diagnostic accuracy of each operator in diagnosing benign, malignant borderline and invasive malignant tumours depending on the level of confidence when suggesting the diagnosis (certain, probable and uncertain).

- Interobserver agreement in classifying tumours as benign, borderline or invasive malignant depending on the degree of diagnostic confidence.

The expert operators' findings were compared to the histology, which was classified according to the World Health Organisation (WHO) guidelines (Serov et al., 1973). Ovarian malignancies were staged according to the classification of the International Federation of Gynaecology and Obstetrics (FIGO) (Shepherd, 1989).
12.3 Statistical Analysis

Please refer to section 11.3.

12.4 Results

The static two-dimensional B-mode images of 171 women diagnosed with adnexal tumours were collected. Five of 171 (3%) cases were excluded by the expert sonologists because of suboptimal image quality.

The 166 cases included in the analysis were assessed by the three expert operators (A, B and C). The data set included 96 (58%) benign ovarian tumours, 34 (20%) borderline ovarian tumours and 36 (22%) primary invasive ovarian tumours. Median age (range) was 57 years (13-88) in patients with a primary invasive ovarian tumour compared to 36 years (21-65) and 39 years (14-83) in women with borderline and benign masses, respectively (p<0.01).

The diagnostic confidence differed significantly between the operators, mainly due to a lower number of probable diagnoses and higher number of certain diagnoses made by operator A than by operators B and C (Table 18). Confidence was also related to the suggested diagnosis in two of the investigators: the rates of “certain” diagnosed masses for operators A, B and C were 44% (15/34), 34% (14/41) and 43% (13/30) in suggested borderline tumours, 60% (58/97), 56% (48/86) and 59% (57/96) in suggested benign tumours and 71% (25/35), 33% (13/39) and 48% (19/40) in suggested primary invasive tumours, respectively (borderline vs. benign A: p=0.11, B: p=0.02, C: p=0.12; borderline vs. invasive A: p=0.02, B: p=0.94 C: p=0.73). The diagnostic performance decreased with decreasing diagnostic confidence (Table 19), e.g. operator A had an error rate of only 6% (6/98) in the ‘certain’ diagnoses, 17% (8/47) in the ‘probable’ diagnoses and 43% (9/21) in the ‘uncertain’ diagnoses (p<0.01). Similar conclusions hold for operators B and C (p=0.01 and p<0.01). Stratified by histology, the figures
suggest similar trends, but statistical significance was only reached in cases of benign tumours (A: \( p=0.05 \), B: \( p=0.01 \), C: \( p<0.01 \)) and in primary invasive tumours for operator A (\( p=0.02 \)). All three investigators misdiagnosed borderline tumours more frequently as benign tumours rather than invasive tumours: A, 89% (8/9); B, 80% (8/10) and C, 86% (12/14).

Table 20 shows that the level of agreement in confidence is proportional to the level of agreement in diagnosis, e.g. operators A and B agreed in confidence in 64% (86/135) of the cases in which they agreed on the diagnosis compared to only 26% (8/31) of the cases in which they disagreed on the diagnosis (\( p<0.01 \)).

Table 21 shows that the interobserver agreement with regard to the diagnoses suggested between any two operators was very good when both suggested a diagnosis with a high level of certainty (agreement rate 98%-100%, Cohen’s Kappa 0.95-1.00), but was significantly lower in case both operators made a probable diagnosis (agreement rate 71%-78%, Cohen’s Kappa 0.64-0.80) (\( p<0.01 \) for any two operators).

The agreement in both diagnosis and confidence was the lowest in the cases of borderline ovarian tumours compared to benign and primary invasive lesions (Table 22).

Table 23 presents the cases which were misdiagnosed by all three investigators. In one case, all three operators were certain that the lesion was benign. The histology report showed evidence of mature cystic teratoma with a focus of squamous cell carcinoma.
Table 18. Diagnostic confidence of the three operators.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Certain n (%)</th>
<th>Probable n (%)</th>
<th>Uncertain N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>98 (59)</td>
<td>47 (28)</td>
<td>21 (13)</td>
</tr>
<tr>
<td>B</td>
<td>75 (45)</td>
<td>81 (49)</td>
<td>10 (6)</td>
</tr>
<tr>
<td>C</td>
<td>89 (54)</td>
<td>70 (42)</td>
<td>7 (4)</td>
</tr>
</tbody>
</table>

Certain: A vs B, p<0.01; A vs C, p=0.19; B vs C, p=0.07
Probable: A vs B, p<0.01; A vs C, p<0.01; B vs C, p=0.18
Uncertain: A vs B, p=0.04; A vs C, p<0.01; B vs C, p=0.47
<table>
<thead>
<tr>
<th>Operator’s confidence in diagnosis</th>
<th>HISTOLOGY</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benign (N=96)</td>
<td>Borderline (N=34)</td>
</tr>
<tr>
<td>Operator A</td>
<td>Incorrectly diagnosed</td>
<td>Correctly diagnosed</td>
</tr>
<tr>
<td>Certain</td>
<td>N=58</td>
<td>3  5%</td>
</tr>
<tr>
<td>Probable</td>
<td>N=28</td>
<td>3  11%</td>
</tr>
<tr>
<td>Uncertain</td>
<td>N=10</td>
<td>3  30%</td>
</tr>
<tr>
<td>Operator B</td>
<td>Incorrectly diagnosed</td>
<td>Correctly diagnosed</td>
</tr>
<tr>
<td>Certain</td>
<td>N=52</td>
<td>5  10%</td>
</tr>
<tr>
<td>Probable</td>
<td>N=41</td>
<td>14  34%</td>
</tr>
<tr>
<td>Uncertain</td>
<td>N=3</td>
<td>1  33%</td>
</tr>
<tr>
<td>Operator C</td>
<td>Incorrectly diagnosed</td>
<td>Correctly diagnosed</td>
</tr>
<tr>
<td>Certain</td>
<td>N=51</td>
<td>-   -</td>
</tr>
<tr>
<td>Probable</td>
<td>N=43</td>
<td>12  28%</td>
</tr>
<tr>
<td>Uncertain</td>
<td>N=2</td>
<td>2  100%</td>
</tr>
</tbody>
</table>
Table 20. Agreement in diagnosis versus agreement in confidence between the operators (N=166).

<table>
<thead>
<tr>
<th>Experts</th>
<th>Agreement in diagnosis (benign, borderline, invasive)</th>
<th>Agreement in confidence (certain, probable, uncertain)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No (N = 31)</td>
<td>74% (N = 49)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Yes (N = 135)</td>
<td>36% (N = 86)</td>
<td></td>
</tr>
<tr>
<td>A, B</td>
<td>No (N = 27)</td>
<td>67% (N = 49)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Yes (N = 139)</td>
<td>35% (N = 90)</td>
<td></td>
</tr>
<tr>
<td>A, C</td>
<td>No (N = 23)</td>
<td>57% (N = 13)</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>Yes (N = 143)</td>
<td>41% (N = 84)</td>
<td></td>
</tr>
<tr>
<td>B, C</td>
<td>No (N = 40)</td>
<td>85% (N = 67)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Yes (N = 126)</td>
<td>53% (N = 59)</td>
<td></td>
</tr>
</tbody>
</table>
Table 21. Interobserver agreement in classifying the tumours as benign, borderline and invasive malignant between each two of the three experts, stratified by agreement in confidence in making the diagnosis (certain and probable).

<table>
<thead>
<tr>
<th>Agreement in confidence</th>
<th>Experts</th>
<th>N</th>
<th>Proportion of agreement in Ultrasound Diagnosis</th>
<th>Kappa (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Benign</td>
<td>Borderline</td>
</tr>
<tr>
<td>Certain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A, B</td>
<td>N=61</td>
<td></td>
<td>42</td>
<td>69%</td>
</tr>
<tr>
<td>A, C</td>
<td>N=70</td>
<td></td>
<td>44</td>
<td>63%</td>
</tr>
<tr>
<td>B, C</td>
<td>N=52</td>
<td></td>
<td>39</td>
<td>75%</td>
</tr>
<tr>
<td>Probable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A, B</td>
<td>N=32</td>
<td></td>
<td>14</td>
<td>44%</td>
</tr>
<tr>
<td>A, C</td>
<td>N=28</td>
<td></td>
<td>13</td>
<td>46%</td>
</tr>
<tr>
<td>B, C</td>
<td>N=42</td>
<td></td>
<td>16</td>
<td>38%</td>
</tr>
</tbody>
</table>

138
Table 22. Rate of agreement in both diagnosis and confidence between the operators, stratified by histology (N = 166).

<table>
<thead>
<tr>
<th>Experts</th>
<th>Histology</th>
<th>Agreement in both diagnosis and confidence</th>
<th>Borderline vs. (benign and invasive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, B</td>
<td>Benign</td>
<td>N = 96, 60</td>
<td>63%</td>
</tr>
<tr>
<td></td>
<td>Borderline</td>
<td>N = 34, 12</td>
<td>35%</td>
</tr>
<tr>
<td></td>
<td>Invasive</td>
<td>N = 36, 14</td>
<td>39%</td>
</tr>
<tr>
<td>A, C</td>
<td>Benign</td>
<td>N = 96, 55</td>
<td>57%</td>
</tr>
<tr>
<td></td>
<td>Borderline</td>
<td>N = 34, 12</td>
<td>35%</td>
</tr>
<tr>
<td></td>
<td>Invasive</td>
<td>N = 36, 23</td>
<td>64%</td>
</tr>
<tr>
<td>B, C</td>
<td>Benign</td>
<td>N = 96, 59</td>
<td>61%</td>
</tr>
<tr>
<td></td>
<td>Borderline</td>
<td>N = 34, 7</td>
<td>21%</td>
</tr>
<tr>
<td></td>
<td>Invasive</td>
<td>N = 36, 18</td>
<td>50%</td>
</tr>
<tr>
<td>A, B, C</td>
<td>Benign</td>
<td>N = 96, 43</td>
<td>45%</td>
</tr>
<tr>
<td></td>
<td>Borderline</td>
<td>N = 34, 5</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Invasive</td>
<td>N = 36, 11</td>
<td>31%</td>
</tr>
</tbody>
</table>
Table 23. Cases misdiagnosed by all three operators.

<table>
<thead>
<tr>
<th>Tumour</th>
<th>Histology</th>
<th>Diagnosis of the experts</th>
<th>Confidence of the experts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td>Cystadenoma</td>
<td>Borderline</td>
<td>Certain</td>
</tr>
<tr>
<td>Benign</td>
<td>Cystadenoma</td>
<td>Borderline</td>
<td>Certain</td>
</tr>
<tr>
<td>Benign</td>
<td>Cystadenofibroma</td>
<td>Borderline</td>
<td>Certain</td>
</tr>
<tr>
<td>Benign</td>
<td>Fibroma</td>
<td>Invasive</td>
<td>Uncertain</td>
</tr>
<tr>
<td>Benign</td>
<td>Fibroma</td>
<td>Invasive</td>
<td>Certain</td>
</tr>
<tr>
<td>Benign</td>
<td>Fibrothecoma</td>
<td>Invasive</td>
<td>Certain</td>
</tr>
<tr>
<td>Borderline</td>
<td>Mucinous</td>
<td>Benign</td>
<td>Certain</td>
</tr>
<tr>
<td>Borderline</td>
<td>Serous</td>
<td>Benign</td>
<td>Certain</td>
</tr>
<tr>
<td>Invasive</td>
<td>Clear Cell carcinoma</td>
<td>Borderline</td>
<td>Certain</td>
</tr>
<tr>
<td>Invasive</td>
<td>with endometriosis</td>
<td>Benign</td>
<td>Certain</td>
</tr>
<tr>
<td>Invvasive</td>
<td>Dermoid with a focus</td>
<td>Benign</td>
<td>Certain</td>
</tr>
<tr>
<td></td>
<td>of squamous ca stage la</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 13

Study 6

Effect of the quality of gynaecological ultrasonography on management of patients with suspected ovarian cancer: a randomised controlled trial

13.1 Background

A clinical diagnosis of adnexal tumours is often made in patients who complain of abdominal swelling and pelvic pain. These tumours can also be noted incidentally during bimanual pelvic examination or on ultrasonography, when an examination is done for early pregnancy complications or other gynaecological symptoms (Beischer et al., 1971; Nelson et al., 1986; Yazbek et al., 2007; Bordgeldt and Andolf, 1999).

The differential diagnosis of adnexal masses includes ovarian cancer and patients are often offered additional investigations, such as tumour marker tests, to clarify the nature of the tumour (Jacobs et al., 1990). It is not unusual for the nature of the adnexal tumour to remain uncertain despite numerous tests and as a result many patients undergo major surgery because of the fear of missing an ovarian cancer (Miller et al., 1991).

Ultrasonography is a sensitive method for the detection of adnexal abnormalities and is routinely used for the assessment of patients with various gynaecological complaints. As previously discussed, the ability of ultrasonography to establish the nature of an adnexal tumour is variable and depends mainly on the experience and skill of the operator (Valentin, 1999; Timmerman et al., 1999b). Studies have shown that, when carried out by experts (known as level III ultrasonography) gynaecological ultrasonography can distinguish between benign and malignant adnexal tumours with an accuracy of 164/173 (94%) tumours (Valentin, 1999). The accuracy of routine ultrasonography (also known as level II ultrasonography), which is usually done by
less experienced operators, is likely to be lower (Timmerman et al., 1999b). The best survival rates for patients with ovarian cancer are achieved when treatment is organised and carried out by gynaecological oncologists who work in cancer centres (Junor et al., 1999). However, patients with asymptomatic benign tumours can be managed expectantly (Yazbek et al., 2007; Platek et al., 1995, Lee et al., 2004) or by minimally invasive surgery if symptomatic (Hilger et al., 2006; Mais et al., 1995). These operations can be safely undertaken by general gynaecologists in local hospitals.

In the UK, most patients with suspected or confirmed adnexal masses are referred to rapid-access gynaecological oncology clinics for detailed investigations and treatment. Most gynaecological ultrasonography examinations, both in local hospitals and tertiary cancer centres are done by sonographers who are usually trained to do level II (EFSUMB Newsletter, 2005) ultrasonography. Before this study, patients who attended our regional gynaecological cancer centre (Guy’s and ST Thomas’ NHS Foundation Trust, London, UK) with an adnexal mass and had non-diagnostic level II ultrasonography, were subsequently referred for Level III scans (EFSUMB Newsletter, 2005) in our unit (King’s College Hospital NHS Foundation Trust, London, UK). Although level III scans were not incorporated in the protocol of the regional gynaecological cancer centre, a retrospective audit showed an improved accuracy of Level III scanning for the diagnosis of ovarian cancers in comparison to level II scans.

The aim of this study was to examine whether the improved accuracy of level III ultrasonography has a measurable effect on the management of patients with adnexal masses when compared with routine level II ultrasonography.

13.2 Methods

We did a prospective randomised controlled study, which took place at Guy’s and St Thomas’ NHS Foundation Trust and at King’s College Hospital NHS Foundation Trust
(London, UK). The study was approved by the ethics committee and the research and development committee at both hospitals.

Patients were recruited from the rapid access clinic at the Southeast Gynaecological Cancer Centre (Guy's and St Thomas' NHS Foundation Trust). The inclusion criteria were: clinical diagnosis of adnexal mass in primary or secondary care or diagnosis of an adnexal mass by ultrasonography, MRI or CT; no indication for immediate surgical treatment and patients aged 14 and above who were able and willing to give written consent to take part in the study.

Randomisation was performed as detailed in section 7.4.

The participants were not blinded to the group assignment. However, the ultrasonography operators were not aware that patients were taking part in the study. Patients who declined taking part in the study were managed in accordance with the cancer centre protocols, which included level II ultrasonography.

Ultrasonography was performed as described in section 7.5. Level II and III operators have also been defined in section 7.5. Level III ultrasonography was done using an Aloka SSD-5000 machine (Aloka Co, Tokyo, Japan) and level II ultrasonography was done using a Siemens Antares™ machine (Siemens Medical Solutions, Bracknell, UK).

The operators were asked to classify tumours as either benign or malignant. When such estimates could not be made, the ultrasonography findings were classified as non-diagnostic and only a morphological description of the tumour was provided to the clinicians. Throughout the study period, level III ultrasonography was done at King's College Hospital NHS Foundation Trust and level II ultrasonography was done at Guy's and St Thomas’ NHS Foundation Trust.

After ultrasonography, patients in both groups of the study were reviewed within two weeks by a gynaecological oncologist at the rapid-access clinic at the Southeast
Gynaecological Cancer Centre (Guy's and St Thomas' NHS Foundation Trust), who reviewed the patient and made decisions about further management. For most patients, the decision on management was made by the lead consultant gynaecological oncologist. In her absence, the decision on management was made by the other two consultant gynaecological oncologists. The gynaecological oncologists were allocated a random mix of the patients as they were recruited into the study. Non-surgical management included expectant management or neoadjuvant chemotherapy when the diagnosis of malignancy was confirmed by cytological or histological examination. If expectant management was chosen, patients attended for follow up ultrasonography within six months of the initial assessment. Patients were reviewed after the follow up scan and surgical management was considered if the patient became symptomatic, the tumour had increased in size or the tumour had changed in appearance (suggesting possible malignancy). Patients with persistent adnexal tumours, which had not changed substantially on follow up visits, were contacted by telephone at least one year after recruitment into the study, and were asked whether they developed any symptoms, whether they needed additional medical help, or whether they had undergone surgery to remove the tumour.

Patients who did not attend follow-up visits were contacted by telephone and asked whether they had undergone ultrasonography in another centre, whether they developed any symptoms, or whether they had undergone a surgical procedure elsewhere.

Surgical management was divided into procedures undertaken for suspected adnexal malignancy or procedures done for presumed benign adnexal tumours. Procedures undertaken for suspected adnexal malignancy involved a major staging surgical procedure (including a laparotomy and at least an oophorectomy and omental biopsy) for a suspected adnexal malignancy at the regional gynaecological cancer
centre. Procedures carried out for benign adnexal tumours included laparotomy by a general gynaecologist or a minimally invasive procedure (eg, operative laparoscopy or ultrasonography-guided cyst aspiration). Patients with suspected non-gynaecological tumours were referred to the appropriate specialty for management.

Findings on ultrasonography were compared with the final histological diagnosis, which was classified according to the World Health Organisation (WHO) guidelines (Serov et al., 1973). Ovarian malignancies were staged according to the classification of the International Federation of Gynecology and Obstetrics (FIGO) (International Federation of Gynecology and Obstetrics, 1971; Shepherd, 1989). Any additional tests used to aid diagnosis were documented. The length of post-operative hospital stay was also recorded in all patients.

The study objective was to examine whether differences in the expertise of ultrasonography operators have a measurable effect on the management of patients with adnexal tumours. The primary endpoint was the number of major staging surgical procedures done in each group. Secondary endpoints include: total number of surgical procedures; number of minimally invasive procedures; number of additional diagnostic tests (eg, CT or laparoscopy); number of follow-up scans; diagnostic accuracy of level II and level III ultrasonography; and duration of hospital stay. The cost-benefit analysis is not part of this paper and will be reported in the future.

A clinical audit performed in the Southeast Gynaecological Cancer Centre between September 2003, and May 2004 showed that 25/55 (45%) patients diagnosed with adnexal pathology underwent a staging laparotomy. Of these, 6/25 (24%) had the diagnosis of ovarian cancer confirmed on histological examination, whereas the remaining 19/25 (76%) patients had benign ovarian disease. Our hypothesis was that level III ultrasonography would halve the number of patients undergoing staging laparotomy. This study was designed to have an 80% power to detect a decrease in the
proportion of staging laparotomies from 45% in the control group to 23% in the study group with a two-sided $\alpha$ of 0.05. The study needed a minimum of 144 patients, but we randomised 150 patients to allow for loss of power due to cancellations or changes in the management plan.

This study is registered on the Current Controlled Trials website http://www.controlled-trials.com/mrct/trial/230201/ISRCTN02631195.

This study was funded by King’s College Hospital NHS Foundation Trust. The funding source had no role in the design, data collection, data analysis, or interpretation of the findings. JY had full access to all the data in the study and had the final responsibility to submit for publication. JY, SR, JBN, TKH, KH and DJ had access to the raw data.

13.3 Statistical Analysis

All statistical analyses were carried out using Medcalc® version 9.2.0.2 (Medcalc Software, Mariakerke, Belgium). Continuous variables were expressed as mean and SD or median and interquartile range. Dichotomous variables were described using proportions. The relative risk (RR) ratio with 95% confidence intervals was used to compare the primary endpoint (number of major staging surgical procedures), the total number of surgical procedures, the number of minimally invasive procedures, the number of follow-up scans, and the number of additional diagnostic tests between the two study groups. Duration of hospital stay after surgery was compared between the two randomised groups using the Mann-Whitney U test. Proportions were compared using Yates corrected $\chi^2$ test. Two-tailed $p$ value of less than 0.05 was considered statistically significant. The diagnostic accuracy of ultrasonography for the diagnosis of ovarian malignancy was assessed by calculating sensitivity, specificity, positive and negative predictive values. The likelihood ratio of a positive test (sensitivity/[1-
specificity]) and the likelihood ratio of a negative test ([1-sensitivity]/specificity) were also calculated in order to eliminate the effect of the prevalence of ovarian malignancy on the interpretation of the value of the diagnostic test (Chien and Khan, 2001).

13.4 Results

The study was done between May 31, 2004 and February 15, 2007. 165 patients met the inclusion criteria, but ten patients declined participation and five were excluded because they needed urgent surgical treatment (Figure 7). The remaining 150 patients were randomised into the study, with 77 randomly assigned to a level III ultrasonography and 73 randomly assigned to a level II ultrasonography. The demographic data of the patients in both groups of the study are presented in Table 24. The two groups were balanced in terms of age, menopausal status, source of referral, and presenting symptoms.

The number of major staging surgical procedures for suspected ovarian malignancy in patients who were randomly assigned to the level II group was significantly higher than the number of major staging surgical procedures in patients randomly assigned to the level III group (27/73 [37%, 95% CI 27 – 48] vs. 17/77 [22%, 95% CI 14 – 33], respectively; RR 1.68 [95% CI 1.00 – 2.81], p=0.049; table 25). The number of planned major staging surgical procedures by gynaecological oncologists in the level II arm was 30/73 (41%, 95% CI 31 - 53) compared with 17/77 (22%, 95% CI 14 - 33) in the level III arm (RR 1.86 [95% CI 1.13 – 3.07], p=0.015). However, three (4%) patients in the level II group declined surgery. One of these three patients eventually had a salpingo-oophorectomy in her local hospital 7 months after randomisation, and the histology showed a stage Ia serous borderline ovarian tumour. Another patient had a follow-up scan after two years, which showed no change in the size or consistency of the tumour, and the third patient was lost to follow-up.
The total number of surgical procedures was 35 (48%, 95% CI 37 – 59) in the level II group, which was not significantly different from 33 (43%, 95% CI 32 – 54) in the level III group (RR 1.12 [95% CI 0.79 – 1.59], p=0.53). 2/73 (3%, 95% CI 1 – 9) patients had minimally invasive procedures in the level II group compared with 9/77 (12%, 95% CI 6 – 21) patients in the level III group (RR 0.23 [95% CI 0.05 – 1.05], p=0.058).

In the subgroup of patients who underwent surgery, a greater number of major staging surgical procedures was done in the level II group compared with the level III group (27/35 [77%, 95% CI 61-88] vs. 17/33 [52%, 95% CI 35-68], respectively; RR 1.50 [95% CI 1.03-2.18]. p=0.036). Minimally invasive procedures were used significantly less in the level II group compared with the level III group (2/35 [6%, 95% CI 2 – 19] vs. 9/33 [27%, 95% CI 15 – 44], respectively; RR 0.21 [95% CI 0.05-0.90], p=0.035).

The median duration of hospital stay for patients who were operated on was 6 days (range 3 – 13) in the level II group and 5 days (range 1 – 9) in the level III group (p=0.01).

The proportion of benign and malignant tumours was not significantly different between the two groups (table 25). In the Level II group, a patient who underwent gynaecological surgery had a right borderline ovarian tumour and a left benign cystadenoma, and was, therefore, classified as having malignant disease. In the level III group, a patient who did not have surgery to remove the adnexal tumour had a biopsy taken during a diagnostic laparoscopy, which showed the presence of a benign uterine leiomyoma. Two patients in the level III group had a CT-guided biopsy, the histology of which showed primary peritoneal cancer, and a patient with intestinal leiomyosarcoma had a laparotomy by a general surgeon. The histological type and stage of the adnexal malignancies are shown in table 26.
10/27 (37%) patients who had a staging laparotomy in the level II arm of the study had malignant disease on histological examination compared with 8/17 (47%) patients in the level III arm (RR 0.78 [95% CI 0.39-1.59], p=0.51). In both groups, no malignant pathology was detected on the histological assessment of patients who were originally planned to have procedures for benign adnexal tumours. However, as previously stated, one patient who declined staging laparotomy was later noted to have a borderline tumour on histological examination.

Initially, 36/73 (49%) patients in the level II group and 41/77 (53%) patients in the level III group were managed expectantly. However, two (3%) of these patients in the level II group and five (6%) of these patients in the level III group were not followed up (RR 0.42 [95% CI 0.09-2.11], p=0.29). Of these, 3/41 (7%) patients in the level III arm of the study died of non-gynaecological causes before attending for the follow-up scan. The remaining four patients (two in each group) were not contactable by telephone and, were thus excluded from the final data analysis. Therefore, outcome data for expectant management were available for 35/73 (48%) patients in the level II group and 36/77 (47%) patients in the level III group (RR 1.03 [95% CI 0.73-1.44], p=0.88; table 27).

The median number of follow-up scans was two (range 0 – 5) in the level II group and one (range 0 – 4) in the level III group (p=0.0004). The mean follow-up period was 427 days (range 1 – 853). Many ovarian cysts resolved spontaneously and only 2/37 (5%) patients managed expectantly in the level II group and 3/41 (7%) patients managed expectantly in the level III group required an intervention (RR 0.44 [95% CI 0.09-2.15], p=0.31). Indications for surgery were: pelvic pain (n=1) and increased size of pelvic tumour on follow-up scan (n=1) in the level II group, and pelvic pain (n=1), persistently high serum CA-125 concentrations (n=1), and suspected
metastatic ovarian tumour (n=1) in the level III group. None of the patients showed malignancy on histology.

A conclusive ultrasonographic diagnosis of the nature of an adnexal tumour was made in 38/73 (52%, 95% CI 41 – 63) patients in the level II group and in 76/77 (99%, 95% CI 93 – 100) in the level III group (RR 0.53, [95% CI 0.42-0.66], p<0.0001). There were two level III and 16 level II ultrasonography operators. The number of conclusive ultrasonography findings among the three level II operators who examined five or more patients were 14/33 scans, 5/10 scans, and 2/5 scans, respectively (p=0.90). The overall number of conclusive scans among the remaining 13 level II operators who examined 1–4 patients was 17/25 scans. The rate of conclusive ultrasonography scans between these two groups of level II operators was not significantly different (p=0.09). The number of non-diagnostic scans did not change significantly over the course of the study. In the first half of the study 22/41 scans were non-diagnostic compared with 13/32 in the second half (p=0.38). The data analysis showed that there were no omission errors in the study, which could have increased the number of non-diagnostic scans.

The accuracy of conclusive diagnosis from ultrasonography was assessed in the subgroup of patients in whom histological findings were also available (table 28). Histological diagnosis was available in 35 (48%) patients in the level II group and in 36 (47%) patients in the level III group. There were three false negative diagnoses of ovarian cancer in the level II group and one false negative diagnosis in the level III group. However, all patients with false negative diagnosis had optimal treatment with a staging laparotomy. The three false negative diagnoses in the level II group were: granulosa cell tumour (stage I), ovarian serous carcinoma (stage III) and a serous borderline ovarian tumour. The false negative diagnosis in the level III group was a mature cystic teratoma containing a focus of stage I squamous cell carcinoma. No false positive diagnoses were noted in the level II group and one false positive diagnosis was
noted in the level III group. This false positive diagnosis was a suspected serous borderline ovarian tumour, which turned out to be a benign cystadenoma.

Serum CA-125 levels were routinely measured in all patients pre-operatively. The median concentrations were not significantly different between the two study groups (19 [range 9-1801] vs. 23 [range 5-2900], p=0.13).
Figure 7 Flow diagram of study participants. US, ultrasound; CT, computed tomography.
Table 24. Demographic data, source of referral and presenting symptoms.

<table>
<thead>
<tr>
<th></th>
<th>Level II</th>
<th>Level III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n= 73</td>
<td>n=77</td>
</tr>
<tr>
<td>Age (years, mean (SD))</td>
<td>48 (17)</td>
<td>49 (18)</td>
</tr>
<tr>
<td>Premenopausal, n</td>
<td>46</td>
<td>43</td>
</tr>
<tr>
<td>Source of referral, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary care</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>Hospital gynaecologist</td>
<td>40</td>
<td>35</td>
</tr>
<tr>
<td>Other specialty</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>Presenting symptoms, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Abdominal swelling</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Other symptoms (PMB, raised CA-125)</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Asymptomatic (incidental finding)</td>
<td>26</td>
<td>31</td>
</tr>
</tbody>
</table>

SD, standard deviation; US, ultrasound; CT, computed tomography; PMB, postmenopausal bleeding.
Table 25. Management of all patients and final histopathological findings from 71 patients.

<table>
<thead>
<tr>
<th>Management, n</th>
<th>Level II, n=73</th>
<th>Level III, n=77</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major staging surgical procedure</td>
<td>27</td>
<td>17</td>
<td>1.68 (1.00-2.81)</td>
</tr>
<tr>
<td>Laparotomy by general gynaecologist</td>
<td>6</td>
<td>7</td>
<td>0.90 (0.32-2.56)</td>
</tr>
<tr>
<td>Minimally invasive procedures</td>
<td>2</td>
<td>9</td>
<td>0.23 (0.05-1.05)</td>
</tr>
<tr>
<td>Laparotomy by other surgeons</td>
<td>0</td>
<td>1</td>
<td>--</td>
</tr>
<tr>
<td>Adjuvant chemotherapy for PPC</td>
<td>0</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Expectant</td>
<td>35</td>
<td>36</td>
<td>1.03 (0.73-1.44)</td>
</tr>
<tr>
<td>Declined surgery</td>
<td>3*</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td>Lost to follow up (expectant)</td>
<td>2</td>
<td>5</td>
<td>0.42 (0.09-2.11)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Histology, n</th>
<th>Level II, n=73</th>
<th>Level III, n=77</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td>25</td>
<td>25</td>
<td>1.05 (0.68-1.64)</td>
</tr>
<tr>
<td>Malignant</td>
<td>10</td>
<td>8</td>
<td>1.31 (0.55-3.16)</td>
</tr>
<tr>
<td>Non-gynaecological cancer</td>
<td>0</td>
<td>3</td>
<td>--</td>
</tr>
</tbody>
</table>

RR, relative risk; PPC, primary peritoneal cancer; US, ultrasound. * One had expectant management, one had surgery at her local hospital and one was lost to follow-up.
Table 26. Histology and International Federation of Gynaecology and Obstetrics stage of malignant adnexal tumours.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Level II ultrasonography (n=10)</th>
<th>Level III ultrasonography (n=8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I, n</td>
<td>I, n</td>
</tr>
<tr>
<td></td>
<td>Invasive epithelial (n=4)</td>
<td>Invasive epithelial (n=3)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Invasive Non-epithelial (n=3)</td>
<td>Invasive epithelial (n=3)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BOT (n=1)</td>
<td>BOT (n=4)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>I, n</td>
<td>I, n</td>
</tr>
<tr>
<td></td>
<td>II, n</td>
<td>II, n</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Invasive epithelial (n=3)</td>
<td>Invasive epithelial (n=4)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Invasive Non-epithelial (n=3)</td>
<td>Invasive epithelial (n=1)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BOT (n=1)</td>
<td>BOT (n=3)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>III, n</td>
<td>III, n</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Invasive epithelial (n=3)</td>
<td>Invasive epithelial (n=4)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Invasive Non-epithelial (n=3)</td>
<td>Invasive epithelial (n=1)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BOT (n=1)</td>
<td>BOT (n=3)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>IV, n</td>
<td>IV, n</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Invasive epithelial (n=3)</td>
<td>Invasive epithelial (n=4)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Invasive Non-epithelial (n=3)</td>
<td>Invasive epithelial (n=1)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BOT (n=1)</td>
<td>BOT (n=3)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

BOT, borderline ovarian tumour.
Table 27. Outcome of the patients managed expectantly.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Level II (n=37)</th>
<th>Level III (n=41)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spontaneous resolution, n</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>Persistent cyst not needing intervention, n</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>Persistent cyst needing intervention, n</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Death from non-gynaecological cause, n</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Lost to follow up, n</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 28. Accuracy of ultrasonography for the diagnosis of malignant adnexal tumours. Data included only patients in whom preoperative diagnosis from ultrasonography was conclusive of nature of the adnexal tumour and for whom histological diagnosis was also available.

<table>
<thead>
<tr>
<th></th>
<th>Level II</th>
<th>Level III</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=15</td>
<td>n=36</td>
<td></td>
</tr>
<tr>
<td>LR+ (95% CI)</td>
<td>$\infty$</td>
<td>24.5 (3.5-170.9)</td>
<td></td>
</tr>
<tr>
<td>LR- (95% CI)</td>
<td>0.60 (0.29-1.2)</td>
<td>0.13 (0.02-0.81)</td>
<td></td>
</tr>
<tr>
<td>Sensitivity, n [% (95% CI)]</td>
<td>2/5 [40 (6.5 - 84.6)]</td>
<td>7/8 [87.5 (47.4 - 97.9)]</td>
<td></td>
</tr>
<tr>
<td>Specificity, n [% (95% CI)]</td>
<td>10/10 [100 (69 - 100)]</td>
<td>27/28 [96.4 (81.6 - 99.4)]</td>
<td></td>
</tr>
<tr>
<td>PPV, n [% (95% CI)]</td>
<td>2/2 [100 (34.2-100)]</td>
<td>7/8 [87.5 (52.9-97.8)]</td>
<td></td>
</tr>
<tr>
<td>NPV, n [% (95% CI)]</td>
<td>10/13 [76.9 (49.7-91.8)]</td>
<td>27/28 [96.4 (82.3-99.4)]</td>
<td></td>
</tr>
<tr>
<td>Area under ROC curve (95% CI)*</td>
<td>0.700 (0.416 - 0.901)</td>
<td>0.920 (0.779 - 0.983)</td>
<td></td>
</tr>
<tr>
<td>Standard error of ROC</td>
<td>0.154</td>
<td>0.069</td>
<td></td>
</tr>
</tbody>
</table>

* p=0.19

LR+, positive likelihood ratio; LR-, negative likelihood ratio; PPV, positive predictive value; NPV, negative predictive value; ROC, receiver-operating characteristics.
Chapter 14

Study 7

Value of preoperative ultrasound in the selection of women with adnexal masses for laparoscopic surgery

14.1 Background

Laparoscopic surgery is the method of choice for treating women with benign adnexal masses. The minimally invasive approach offers many advantages over laparotomy, including decreased postoperative pain, faster recovery and reduction in postoperative intra-abdominal adhesions (Lin et al., 1995; Howard, 1995; Mais et al., 1995). As a result, many operations can be performed as day-surgery procedures, with patients being discharged home a few hours after operation. In order to ensure optimal use of day surgery facilities it is important to ensure that the risk of converting a laparoscopic procedure to a laparotomy is minimal. The factors that may increase the risk of laparotomy have not been defined scientifically, but there is a general consensus that severe pelvic adhesions and pelvic endometriosis can cause difficulties during pelvic surgery and increase the risk of complications. It is also known that spillage of material from dermoid cysts can occasionally cause severe peritonitis and is best avoided (Huus et al., 1996; Coccia et al., 1996; Remorgida et al., 1998). Therefore some authors advocate open surgery for large dermoid cysts >10 cm in mean diameter as they are difficult to remove form the abdominal cavity without rupture (Mettler et al., 2001). Finally, laparoscopic surgery should not be performed on women with invasive ovarian cancer (Parker, 1992; Kindermann et al., 1995). There is a consensus in the UK that optimal care for these women involves a laparotomy by a gynaecological oncologist in a regional cancer centre (Junor et al., 1999).
Ultrasound examination is the main tool used to diagnose adnexal tumours and to differentiate between benign and malignant lesions. The value of ultrasound and Doppler in correctly identifying the risk of malignancy of adnexal tumours and selecting those with low risk of malignancy for laparoscopic surgery has been assessed previously (Berlanda et al., 2002; Guerriero et al., 2005). However, the value of ultrasound for the preoperative selection of women with benign adnexal masses for removal by either laparoscopic surgery or by laparotomy has not been examined.

The aim of this study was to prospectively develop and assess a set of ultrasound criteria that were designed to select women for laparoscopic removal of ovarian cyst with a minimum risk of converting the operation to a laparotomy.

14.2 Methods
This prospective observational study was carried out at the Gynaecological Assessment Unit at King's College Hospital. Symptomatic women (abdominal pain, increased abdominal girth or abdominal mass) with a clinical or ultrasound diagnosis of an adnexal mass were offered a detailed transvaginal scan in order to assess the feasibility of laparoscopic surgery. All women were assessed by gynaecologists trained in transvaginal ultrasound. The clinical history was recorded and a brief clinical examination was performed. All women then underwent a gynaecological ultrasound examination using transvaginal and transabdominal transducers with B-mode and colour Doppler facilities (Aloka SSD-2000 or Aloka SSD-5000, Aloka Co, Tokyo, Japan).

Ultrasoundography was carried out as detailed in section 7.2.

Severe endometriosis was suspected in women with evidence of ovarian endometriomas in whom the ovaries were firmly adherent to the posterolateral aspect of the uterus or in whom the pouch of Douglas was obliterated with adhesions (Ghezzi et al., 2005; Bhatt et al., 2006). The criteria we used to diagnose obliterated pouch of
Douglas with adhesions were: the absence of free movement between the posterior uterine surface and the large bowel.

Pelvic adhesions were suspected when the adnexal tumour could not be mobilised by using gentle palpation using the transvaginal or transabdominal probe or in women with evidence of peritoneal pseudocysts.

All postmenopausal women and those older than 45 were booked for oophorectomies, whereas the remaining women were offered cystectomies.

Women with presumed benign lesions who did not meet the selection criteria for laparoscopic surgery were booked for a laparotomy. In addition, patients with adnexal cysts underwent a laparotomy, if they also required myomectomy or hysterectomy. All women with suspected cancers were referred for further management to the Regional Cancer Centre. These women were not included in our study.

Operation details were recorded in all patients. All the operations were carried out by a single team of surgeons, all of whom had considerable experience in intermediate-level laparoscopic surgery (RCOG in collaboration with BSGE, 2005). An intermediate-level laparoscopic surgeon would be able to independently perform laparoscopic procedures such as the treatment of ectopic pregnancy, ovarian cystectomy, oophorectomy, and excision or ablation of peritoneal endometriosis and endometriomas.

Laparoscopic surgery was classified as successful if the adnexal mass was removed completely without resorting to laparotomy. Laparoscopic ovarian cystectomy was classified as successful when the operation was completed with the conservation of the ipsilateral ovary.

The ultrasound findings were compared to the histology, which was classified according to the World Health Organisation (WHO) guidelines (Serov et al., 1973).
14.3 Statistical Analysis

All statistical analyses were carried out using Medcalc® version 9.2.0.2 (Medcalc Software, Mariakerke, Belgium). The statistical significance of differences in continuous variables was determined using Mann-Whitney U-test, Kruskall-Wallis test or Student t-test depending on data distribution. Proportions were compared using Yates’-corrected Chi-square test. Two-tailed p <0.05 was considered statistically significant. The diagnostic accuracy of the tests was assessed using sensitivity, specificity, positive predictive value (PPV) and positive (LR+) and negative (LR-) likelihood ratio measures.

14.4 Results

A total 143 women with 162 benign adnexal masses on ultrasound scan were scheduled for surgery (Figure 8). Based on ultrasound assessment, 119 women were considered suitable for laparoscopic surgery. Two (1.7%) women also required myomectomy and one (0.8%) woman needed a hysterectomy and they were booked for a laparotomy instead. Laparotomy was also performed in three women (2.5%) who refused laparoscopy. These six (5%) women were excluded from final data analysis.

The final dataset consisted of 137 women, 113 (82.5%) of whom were booked for a laparoscopy and 24 (17.5%) for a laparotomy. The indications for laparotomy were: dermoid cysts >10 cm in mean diameter in 10 (41.7%) women, suspected severe pelvic adhesions in six (25%), solid ovarian tumours > 5 cm in mean diameter in five (20.8%) and severe pelvic endometriosis in three (12.5%).

There were no statistically significant differences between women selected for laparoscopy and those selected for laparotomy in regard to their age, menopausal status, parity, history of previous laparoscopic surgery, history of previous laparotomy and history of previous ovarian surgery (Table 29). However, there was a statistically
significant difference in the median tumour volume between the two groups (Table 29). A higher proportion of unilocular cysts were operated on laparoscopically, and a higher proportion of multilocular solid cysts were operated on by laparotomy (p<0.05).

The indication for laparotomy was confirmed at operation in 22/24 (91.7%; 95% CI, 74–98) women who were booked for open surgery. In the remaining two (8.3%; 95% CI, 2-26) cases the ultrasound diagnosis of severe adhesions was not confirmed at surgery.

Some 113/137 (82.5%; 95% CI, 75-88) women were considered suitable for laparoscopic surgery. Of these, 88 (77.9%; 95% CI, 69-85) were booked for laparoscopic ovarian cystectomy and 25 (22.1%; 95% CI, 15-31) were booked for laparoscopic oophorectomy.

In all, 107/113 (94.7%; 95% CI, 89-98) women had their operation completed laparoscopically, and in the remaining six (5.3%; 95% CI, 2-11) women, the operation was converted into a laparotomy. The reasons for failed laparoscopy were: extensive pelvic adhesions in three (50%; 95% CI, 19-81) cases, morbid obesity in two (33.3%; 95% CI, 10-70) cases and injury to the inferior epigastric artery in one (16.7%; 95% CI, 3-56) case. There was no significant difference in the median volume of ovarian cysts of women who had a successful or failed laparoscopy. There was also no association between failed laparoscopy and the presence of bilateral ovarian cysts (p>0.05).

Laparoscopic ovarian cystectomy was successful in 81/88 (92%; 95% CI, 84-96) cases and the operation was converted to laparoscopic oophorectomy in the remaining seven (8%; 95% CI, 4-15) women. The reasons for conversion to laparoscopic oophorectomy were: heavy bleeding from the cyst site in four (57.1%; 95% CI, 25-84) cases, severe periovarian adhesions in two (28.6%; 95% CI, 8-64) cases and in one (14.3%; 95% CI, 3-51) case the adnexa appeared necrotic owing to previous torsion. The median tumour volume in cases of successful laparoscopic ovarian cystectomy was...
67 (range 4–1972) mL, which was significantly less in comparison to 229 (range 29–1231) mL in failed laparoscopic ovarian cystectomy (Z=1.97, p<0.05).

The preoperative ultrasound assessment predicted the outcome of laparoscopic surgery with a sensitivity of 98% (95% CI, 0.94–0.99), specificity of 79% (95% CI, 0.60–0.90), PPV of 95% (95% CI, 0.89–0.98), LR+ of 4.58 (95% CI, 2.25–9.32) and LR- of 0.02 (95% CI, 0.01–0.09). Preoperative ultrasound assessment also predicted the success of laparoscopic ovarian cystectomy with specificity of 78% (95% CI, 0.61–0.89), PPV of 92% (95% CI, 0.85–0.96) and LR+ of 4.6 (95% CI, 2.38–8.80).

On histological examination, the diagnosis of benign adnexal cyst was confirmed in 152/153 (99.3%) cysts. All bilateral tumours in this study were of same histological type. One woman who had a laparoscopy for a presumed benign ovarian cyst was diagnosed with a gastrointestinal type borderline ovarian tumour. There were no cases of invasive ovarian cancer in either group of patients.

There was a statistically significant difference between histological subtypes in women who had laparoscopic surgery and those who had laparotomy (χ² = 35.61, P<0.001) (Table 30). Histological tumour type was correctly predicted in 141/153 (92.2%) cysts. The false positive cases were: four (33.4%) endometriomas (one misdiagnosed as dermoid cyst and three as cystadenomas), two (16.7%) dermoid (misdiagnosed as cystadenomas), one (8.3%) cystadenoma (misdiagnosed as endometrioma), one (8.3%) thecoma (misdiagnosed as endometrioma), one (8.3%) mucinous gastrointestinal type borderline ovarian tumour (misdiagnosed as cystadenoma), one (8.3%) pseudocyst (misdiagnosed as fimbrial cyst), one (8.3%) simple ovarian cyst (misdiagnosed as fimbrial cyst) and one (8.3%) hydrosalpinx (misdiagnosed as cystadenoma). These cysts shared a common morphological appearance, which was a unilocular cyst with echogenic fluid content.
There were no dermoid cysts >10cm found at laparoscopy. Nor were there any cases of severe endometriosis encountered in the laparoscopy group. In three cases where severe endometriosis was suspected, laparotomy was performed and the ultrasound findings were confirmed at operation.

Five cases of severe adhesions were missed on ultrasound in the laparoscopy group and there were two false positive diagnoses of adhesions in the laparotomy group. The preoperative diagnosis of severe pelvic adhesions was made with a sensitivity of 44% (95% CI, 17–69), specificity of 98% (95% CI, 94–99), PPV of 67% (95% CI, 30–90), LR+ of 28.4 (95% CI, 6.00 – 134.96) and LR- of 0.56 (95% CI, 0.31–1.01).
143 women scheduled for surgery

119 women suitable for laparoscopic surgery

24 women not suitable for laparoscopic surgery

6 women excluded:
- 2 required myomectomy
- 1 required hysterectomy
- 3 declined laparoscopic surgery

Laparotomy

88 women booked for ovarian cystectomy

81 successful

7 converted to laparoscopic oophorectomy:
- 4 heavy bleeding
- 2 adhesions
- 1 previous torsion

25 women booked for oophorectomy

19 successful

6 converted to laparotomy:
- 3 adhesions
- 2 obesity
- 1 vascular injury

Figure 8 Flow chart of the patients in the study.
Table 29. Demographic features of women and adnexal tumour characteristics of women booked for laparotomy or laparoscopy.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Laparoscopy</th>
<th>Laparotomy</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention to treat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years, median (range))</td>
<td>35 (14-72)</td>
<td>32 (20-62)</td>
<td>NS</td>
</tr>
<tr>
<td>Premenopausal (n (%))</td>
<td>99 (88)</td>
<td>22 (92)</td>
<td>NS</td>
</tr>
<tr>
<td>Parity (median (range))</td>
<td>1 (0-4)</td>
<td>0 (0-4)</td>
<td>NS</td>
</tr>
<tr>
<td>Previous laparoscopy (n (%))</td>
<td>13 (12)</td>
<td>3 (13)</td>
<td>NS</td>
</tr>
<tr>
<td>Previous laparotomy (n (%))</td>
<td>23 (20)</td>
<td>9 (38)</td>
<td>NS</td>
</tr>
<tr>
<td>Ultrasound tumour appearance* (n (%))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unilocular</td>
<td>61 (54)</td>
<td>4 (17)</td>
<td>0.002</td>
</tr>
<tr>
<td>Unilocular solid</td>
<td>8 (7)</td>
<td>4 (17)</td>
<td>NS</td>
</tr>
<tr>
<td>Multilocular</td>
<td>27 (24)</td>
<td>5 (21)</td>
<td>NS</td>
</tr>
<tr>
<td>Multilocular solid</td>
<td>7 (6)</td>
<td>6 (25)</td>
<td>0.01</td>
</tr>
<tr>
<td>Solid</td>
<td>1 (1)</td>
<td>2 (8)</td>
<td>NS</td>
</tr>
<tr>
<td>Not defined</td>
<td>9 (8)</td>
<td>3 (12)</td>
<td>NS</td>
</tr>
<tr>
<td>Previous ovarian surgery (n (%) )</td>
<td>6 (5)</td>
<td>3 (13)</td>
<td>NS</td>
</tr>
<tr>
<td>Mass volume (mL, median (range))</td>
<td>70 (1–1972)</td>
<td>242 (0.3–2907)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

*All bilateral ovarian tumours in this study had similar ultrasound appearance. NS, not significant.
Table 30. Histological subtypes of adnexal tumours operated on by laparoscopy and laparotomy.

<table>
<thead>
<tr>
<th>Histological subtype</th>
<th>Operated by</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Laparoscopy</td>
<td>Laparotomy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dermoid</td>
<td>55 (44.4)</td>
<td>18 (62)</td>
<td></td>
<td>73</td>
</tr>
<tr>
<td>Cystadenoma</td>
<td>33 (26.6)</td>
<td>1 (3.5)</td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>Endometrioma</td>
<td>5 (4.0)</td>
<td>3 (10.3)</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Fibroma</td>
<td>0 (0.0)</td>
<td>5 (17.2)</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Thecoma</td>
<td>1 (0.8)</td>
<td>0 (0.0)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Borderline</td>
<td>1 (0.8)</td>
<td>0 (0.0)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Simple cyst</td>
<td>19 (15.3)</td>
<td>0 (0.0)</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>Hydrosalpinx</td>
<td>7 (5.7)</td>
<td>1 (3.5)</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Pseudocyst</td>
<td>3 (2.4)</td>
<td>1 (3.5)</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>124 (100)</td>
<td>29 (100)</td>
<td></td>
<td>153</td>
</tr>
</tbody>
</table>
PART THREE:

Discussions
Chapter 15

Study 1

A comparative study of the risk of malignancy index and the ovarian crescent sign for the diagnosis of invasive ovarian cancer

Our study confirms previous reports (Jacobs et al., 1990; Hillaby et al., 2004), which showed that RMI and OCS are fairly good methods for preoperative differentiation between invasive and non-invasive adnexal tumours. Both tests have proven to be very sensitive for the detection of invasive ovarian cancer. However, their false positive rates were well above 5%. False positive rates are particularly important when the tests are applied to low risk population diagnosed with ovarian tumours during opportunistic screening for ovarian abnormalities.

The RMI was particularly sensitive to elevations of serum CA-125 which was the main reason for high readings in benign pathology, 87.5% of which had ovarian endometriomas. The OCS was typically absent in larger tumours such as cystadenomas and borderline ovarian tumours, but it was always easily detectable in endometriotic cysts. Neither the RMI nor the OCS was accurate enough in the diagnosis of borderline ovarian tumours. The RMI failed to diagnose all five borderline ovarian tumours in our study population, whilst the sensitivity of the ovarian crescent sign was only 60%. This is an expected result as borderline ovarian tumours exhibit different morphological and biochemical characteristics to invasive cancers and therefore they are unlikely to be detectable using tests, which are primarily designed to detect invasive disease. There is some evidence that pattern recognition method may help to correctly diagnose some borderline tumours (Alfuhaid et al., 2003; Exacoustos et al., 2005; Fruscella et al., 2005), but this approach has not been prospectively tested as yet.

The RMI has been used in clinical practice for many years, while the OCS is a novel diagnostic test, which has been used by a limited number of units. One of the
important advantages of RMI is its relative simplicity in the approach to ultrasound assessment of ovarian tumours. Rather than relying on complex ultrasound tests such as Doppler assessment of intraovarian blood flow or on the meticulous analysis of minute details of tumour architecture, The RMI simply divides ovarian cysts into unilocular and complex. This approach requires minimal ultrasound skills and it may be successfully applied in units without high level of expertise in gynaecological ultrasonography.

Many of the described ultrasound-based diagnostic models are too complex for the use in daily practice and they may not be easily transportable between different ultrasound units (Tailor et al., 1997; Timmerman et al., 1999a; Valentin et al., 2001; Aslam et al., 2000a). The OCS overcomes this problem, and similarly to the RMI, it only requires a modest level of expertise in gynaecological ultrasonography. It is therefore likely that the OCS may be successfully used in units with average ultrasound skills who are currently relying on RMI.

Our study may be criticised for failing to obtain histological confirmation of the nature of ovarian tumours in all the cases. However, in modern clinical practice it is not acceptable to operate on all asymptomatic women with an incidental ultrasound diagnosis of a benign looking cyst. Nevertheless, all women in our study had follow up examinations and non-surgical treatment was offered only to women with cysts, which were not increasing in size.

Our results show that false positive results with the RMI and OCS occur in different types of ovarian tumours. Therefore, a combination of the two tests may be an effective way of reducing false positive findings.

This study could also be criticised for including a small number of invasive ovarian cancers especially stage I cancers as well as a small number of borderline ovarian tumours. However, women who were included in our study attended for an ultrasound scan not exclusively because of a known ovarian tumour and only a study
performed at a regional cancer centre over a long period of time would pick up a higher rate of invasive ovarian malignancies or borderline ovarian tumours.

Future work carried out within a gynaecological cancer centre including a high number of women with borderline and invasive ovarian tumours will show whether the OCS may be a more effective way, compared to the RMI, of managing women with ultrasonographic diagnosis of ovarian tumours.
Chapter 16

Study 2

The accuracy of ultrasound subjective “pattern recognition” for the diagnosis of borderline ovarian tumours

Our study has confirmed that borderline ovarian tumours tend to occur in younger women and that they are detected incidentally during clinical or ultrasound examination in the majority of cases for indications other than a suspected ovarian tumour. The differences in age and clinical symptoms between borderline and invasive ovarian tumours were statistically significant and should be taken into consideration when assessing an adnexal tumour by ultrasound. These observations were in agreement with the results of previous studies, which also showed that women diagnosed with a BOT were more likely to be asymptomatic than those with benign or invasive ovarian tumours (Gotlieb et al., 2005; Boran et al., 2005; Webb et al., 2004; Gotlieb et al., 1998).

We were able to establish a correct histological diagnosis in more than two-thirds of all BOTs. The specificity of ultrasound diagnosis was very high, but the sensitivity was only 69%. This was mostly due to the relatively high proportion of cysts, which did not exhibit typical ultrasonic features of BOTs. Amongst 11 false-negative cases encountered in our study, there were four (36%) cysts, which were unilocular with smooth inner walls and no papillary projections. In all these four cases, there was a complete agreement between the ultrasound morphological appearances and the histopathological macroscopic description in misdiagnosed ovarian tumours. There were no cases in which papillary projections were missed on ultrasound scan when the results were compared to histology. Nonetheless, our data do not support surgical intervention in all cases of unilocular cysts, although it would be fair to say that the
diagnosis of BOT needs to be considered in women with large and growing unilocular cysts.

The proportion of unilocular cysts in our population of women with BOT was 11.4%, which was much higher than the findings of a retrospective study by Fruscella and co-workers (2005), who found unilocular cysts in only 3.5% of their population of BOTs.

However, we were able to confirm their observation that different subtypes of BOTs display different morphological features, which is helpful in the differential diagnosis. We also agree with Fruscella and co-workers’ (2005) findings that serous / endocervical type BOTs are characterised by a higher rate of unilocular solid lesions, higher number of papillary projections and lower prevalence of multilocular lesions when compared with GI type BOTs. Both studies showed a statistically significant difference between the morphological ultrasound appearances of serous / endocervical type and GI type BOT. We were more accurate in diagnosing serous / endocervical type BOT compared to GI type BOT. This is important in view of the relatively worse prognosis of serous / endocervical type BOT and the higher recurrence rate (Ahmed and Lawton, 2005).

The presence of a honeycomb nodule was highly specific for GI type BOTs. Although, slightly more than half GI type BOT had a honeycomb nodule, this nodule was present in a case of benign mucinous cystadenoma. On the other hand, all GI type BOTs had thick fluid noted within the cyst, which makes this finding a highly sensitive marker for this subtype of mucinous BOTs. Thick fluid content was also a characteristic of the majority of ovarian endometriomas, a few serous / endocervical type BOTs and some epithelial malignant ovarian tumours.

Our description of GI type BOTs is not substantially different from the one provided by Fruscella and co-workers (2005) who described the typical pattern of
mucinous gastrointestinal type BOT as being a multilocular cyst with a high number of locules. Using International Ovarian Tumor Analysis (IOTA) (Timmerman et al., 2000) group classification these tumours should indeed be described as multilocular cysts. The honeycomb nodule, however, is a specific feature of GI type BOT, which represents a subgroup of multilocular cysts. This distinction is important, as there are many benign and invasive tumours, which are classified as multilocular cysts. Nearly two-thirds of benign cystadenomas in this study were described as multilocular cysts and, without performing a more detailed analysis, they could have all been wrongly classified as GI type BOTs.

The ovarian crescent sign was present in 75% of serous / endocervical type BOTs, but only in 20% of GI type BOTs. This was mainly due to the large size of GI type BOT, which makes the visualisation of healthy ovarian tissue more difficult. However, in the presence of a positive ovarian crescent sign, the diagnosis of an invasive tumour is very unlikely. There was only one case of an invasive tumour with a positive crescent sign in this series, which was a mature cystic teratoma with a focus of squamous cell carcinoma (Stage I). In practical terms the ovarian crescent sign can be used to exclude an invasive ovarian cancer, but its absence is not diagnostic of invasive ovarian disease.

In this study, the overall sensitivity of pattern recognition in differentiating between benign and malignant tumours (including both borderline and invasive lesions) was 91% (95% CI, 83 – 95), which was similar to the findings by Valentin et al. (2000) who reported a sensitivity of 83% (95% CI, 67 – 94). However, the specificity in our study was less (85% (95% CI, 75 – 92) vs. 91% (95% CI, 84 – 96)), which may be explained by a higher proportion of BOTs in our study population.

Our data show that the preoperative ultrasound diagnosis of BOT using the pattern recognition method is highly specific. Thus, women who are in the reproductive
age group could benefit from fertility sparing surgery in a regional gynaecological cancer centre. On the other hand, when reproductive potential is not an issue, accuracy in preoperative diagnosis would allow better preoperative counselling and a more conservative surgical approach.

All the correctly diagnosed BOTs and one false negative case (thought to be an invasive cancer) underwent surgical staging. Staging remains important in cases of serous / endocervical type BOTs, due to the possibility of advanced stage disease at diagnosis (Silverberg et al., 2004), especially in the case of micropapillary subtype serous BOT. The importance of staging may be less in GI type BOT and the impact of a complete staging on prognosis of BOTs is far from clear (Gotlieb et al., 1998; Fauvet et al., 2004).

A relatively high number of women in this study had early stage borderline and invasive ovarian tumours. However, this finding is not representative of all the women who were operated on at the regional gynaecological cancer centre as our study only included the proportion of women in whom the nature of the lesion was indeterminate during the level II scan. Women with conclusive features of ovarian cancer or benign tumours were managed according to the centre’s protocols as stated in the methods and these were not included in our study.

It is also unclear whether spillage of the cyst content adversely affects the prognosis of women with BOTs. This is particularly an important issue in the context of laparoscopic surgery for ovarian tumours. Laparoscopic surgery offers many advantages to women with benign cysts because of the shorter hospital stay, reduced postoperative pain, faster recovery and return to normal daily activities (Lin et al., 1995; Howard, 1995; Mais et al., 1995). However, the risk of cyst spillage during laparoscopy may be much greater in comparison to open surgery. This opinion is not shared by Fauvet et al. (2005) who think that the reason for ovarian cyst rupture was not related to
the surgical route but to the frequency of cystectomy. No wound metastasis occurred in their series (Fauvet et al., 2005). Our data show, that even in the optimal circumstances a number of BOTs will be misclassified as benign on ultrasound scan and therefore considered to be suitable for minimally invasive surgery. Thus, there is a great need to continue improving non-invasive diagnosis of borderline tumours and, at the same time, to assess the effects, if any, of BOT spillage on the survival and recurrence rates (Maneo et al., 2004; Fauvet et al., 2005). Until such data are available, every effort should be made to minimise the risk of spillage of cyst content during surgery for all types of ovarian cysts.

This study only assessed the value of pattern recognition in diagnosing BOTs and differentiating them from benign and invasive ovarian tumours. Doppler studies were not used in the differentiation between different types of ovarian tumours in this study. To our knowledge, no studies have investigated the role of this modality in the diagnosis of BOTs and, perhaps, the value of Doppler studies in aiding the diagnosis of BOTs should be evaluated in a prospective randomised study.

Our dataset included a significant number of functional and dermoid cysts, which are usually simple to diagnose on a routine ultrasound examination. They were often referred to the regional cancer centre because of raised serum CA-125 levels in association with the cyst. In such situations it is not unusual for gynaecological oncologists to refer these patients for an expert scan even if non-specialist ultrasound findings were reassuring.

In conclusion, our study has shown that the correct ultrasound diagnosis of BOT can be achieved in two-thirds of cases. Further work is required to identify morphological and biochemical features that may help to increase the accuracy of pre-operative detection of such tumours.
Chapter 17

Study 3

Real-time ultrasound versus evaluation of static images in the preoperative assessment of adnexal masses

Our results indicate that predicting the character of an adnexal mass by evaluating static ultrasound images is less accurate than by evaluating ultrasound findings during a real-time scan. The difference in accuracy is explained by a difference in specificity, and it seems that ultrasound examiners tend to be more confident in excluding malignancy during a real-time scan than when looking at saved static ultrasound images of an adnexal mass. Because a second opinion is more frequently requested in difficult cases than in obviously benign or malignant cases, it was deliberate that our study population comprised mainly 'difficult' and complex masses, i.e., an unrealistically high number of borderline tumours, which are known to be difficult to classify as benign or malignant using pattern recognition (Valentin et al., 2006). Masses that can be obviously identified as being benign and malignant on real-time ultrasound examination can probably also be more easily and accurately classified (in comparison to difficult cases) by the assessment of static ultrasound images, and in such cases the difference in accuracy between real-time and static image assessment that was found in this study may not exist.

It is likely that the performance of the image experts in this study would have been poorer if the real-time examination had been performed, and the static ultrasound images saved, by a less experienced sonologist. An expert sonologist is more likely to be able to create representative ultrasound images of an adnexal mass demonstrating relevant ultrasound features. On the other hand, it is possible that the performance of the image experts would have been better if they had had the possibility to evaluate a
volume of static images collected by three-dimensional ultrasound instead of only a few selected two-dimensional ultrasound images. However, when looking at static images, the dynamic aspect of the real-time ultrasound examination is lost. On static images a liquefying blood clot in a haemorrhagic cyst or a recent bleeding in an endometrioma may be misinterpreted as an irregular papillary projection of solid tissue, but during a real-time ultrasound examination, clots may be seen sliding against the cyst wall when the cyst is pushed on with the ultrasound probe. Blood clots mimicking a multilocular cyst on a static image may demonstrate the typical “jelly” movement when pushed on during a real-time scan. On a static image, adhesions may give the impression of thick septa in a cystic tumour, while gently pushing on them with the probe will show the “flapping sail sign” (Savelli et al., 2004). Thus, the performance of the image experts might have been better if they had been presented with representative video clips or with the use of four-dimensional cine loops of volume instead of with static images. Failure to use video clips in our study was a missed opportunity and this should be investigated in a future work.

In conclusion, a preoperative diagnosis of an adnexal mass made on the basis of a real-time ultrasound examination is more accurate than a diagnosis made on the basis of saved static ultrasound images. Evaluation of static images is associated with lower specificity. The dynamic aspect of the real-time scan gives the examiner supplementary information and makes it possible to scrutinise every part of the mass.
Chapter 18

Study 4

The use of ultrasound pattern recognition by expert ultrasound operators to identify borderline ovarian tumours: a study of diagnostic performance and interobserver agreement of ultrasound diagnoses

This study has confirmed previous findings that ultrasound pattern recognition method is a reproducible and accurate test for differential diagnosis of adnexal tumours (Valentin, 1999; Timmerman et al., 1999b). The level of interobserver agreement ranged from substantial to very good, when the operators were asked to assess the nature of adnexal lesions by ultrasound. This study has also shown that borderline ovarian tumours are more difficult to diagnose correctly than benign and invasive malignant ovarian tumours. The decline in diagnostic performance in cases of borderline ovarian tumour was observed in all three expert operators.

The diagnostic performance of the expert sonologists was higher when discriminating invasive from non-invasive (benign and borderline) ovarian tumours than in discriminating malignant (borderline and invasive) from benign masses, mainly due to tendency to misclassify borderline ovarian tumours as benign lesions. As the interobserver agreement in the former was also better, these findings suggest that the outcomes could be dichotomised into non-invasive and invasive malignant lesions, as opposed to traditional diagnosis of benign and malignant lesions.

These findings were in agreement with our second study, which showed a low sensitivity of ultrasound pattern recognition method for the diagnosis of borderline ovarian tumours with 11% (4/35) of them presenting as simple cysts, which could not be differentiated from benign cysts. This problem was mostly encountered in cases of mucinous endocervical and serous types borderline ovarian tumours, whilst the
diagnostic accuracy was better in cases of mucinous gastro-intestinal type borderline ovarian tumours.

Provided the results of our study are confirmed by others, ultrasound operators using pattern recognition method may be advised in the future to describe adnexal lesions as invasive or non-invasive, rather than as benign or malignant. The drawback of this approach is that a proportion of women with borderline lesions may be offered minimally invasive surgery, which increases the risk of the spillage of tumour content into the abdominal cavity. There is some evidence in the literature that spillage of invasive epithelial ovarian tumours at surgery may adversely affect patient's long term outcomes (Vergote et al., 2001). The prognosis of borderline tumours, particularly those of gastro-intestinal type is very good and there is no strong data to suggest that the spillage of cyst fluid in these cases could cause harm (Fauvet et al., 2005). Nevertheless, it is prudent to take measures to minimise the risk of spillage until more evidence is available about its effect on women’s health.

Our previous work also showed that the specificity of ultrasound diagnosis of borderline tumours is high. Therefore the risk of spillage may be minimised if ultrasound operators are encouraged to report the diagnosis of a borderline ovarian tumour whenever typical morphological features are present on ultrasound scan (Exacoustos et al., 2005; Fruscella et al., 2005).

Our study may be criticised for being performed on images stored on our database, rather than being based on real-time examination. However, the expert operators were given the option of removing cases from the study if they felt that there was insufficient data to form an opinion about the nature of an adnexal mass. In more than 97% (166/171) of the cases they felt that the information provided was satisfactory, which is probably not different from live scanning. It is also important to remember that
evaluation of static ultrasound images by radiologists is standard practice in many units around the world.

The relative proportion of borderline ovarian tumours in our dataset was higher than one would expect to see in routine clinical practice, which could have had a negative effect on the diagnostic accuracy and reproducibility. The operators were aware, however, that the main objective of the study was to assess their accuracy in detecting borderline ovarian tumours, which could have improved their ability to diagnose these tumours.

In conclusion, our study has confirmed that ultrasound pattern recognition is an accurate and highly reproducible method for characterisation of adnexal tumours. The test, however, performs less well in cases of borderline ovarian tumours, which are sometimes difficult to distinguish from benign tumours. If this finding is reproduced by others, the structure of reporting ultrasound findings may need to change in order to classify ovarian tumours as being non-invasive or invasive, rather than making a diagnosis of benign or malignant tumour.
Chapter 19

Study 5

Confidence of the ultrasound operator in making a differential diagnosis of an adnexal tumour: Effect on diagnostic accuracy and interobserver agreement

This study showed a clear association between the level of confidence and the diagnostic performance as the diagnostic accuracy of all three operators decreased in proportion to decreasing diagnostic confidence. Nonetheless, the three ultrasound operators, despite their similar high expertise, had variable degrees of confidence when they suggested a diagnosis of benign, borderline or invasive malignant ovarian tumours. Pattern recognition was a very useful test for discriminating between benign, borderline and invasive malignant ovarian tumours if the operators made their diagnoses with certainty, but if the operators were uncertain about their diagnoses, it was a virtually useless test.

Diagnostic certainty is usually associated with knowledge in medical decision making. ‘Assessing confidence and its individual components may prove extremely difficult as this may be influenced by complex factors, which interact and influence the level of confidence. These include the amount of training or expertise, time allocated to decision making, the degree of self-confidence, cultural differences and the operator’s internal mental process influenced by memory and emotions, which interact with the external environment and problem context’ (Rohrbaugh and Shanteau, 1999; Baranski and Petrusic, 1998).

The agreement in both diagnosis and confidence, in our study, was the lowest in the case of borderline ovarian tumours. Borderline ovarian masses were most often misclassified as benign tumours. This could be because certain types of benign and borderline ovarian tumours share similar morphological features, which make the
differentiation between these two types of tumours difficult (Exacoustos et al., 2005; Fruscella et al., 2005).

The inter-observer agreement was almost perfect when the operators were certain about their final diagnoses (Cohen’s Kappa 0.95-1.00). This high level of agreement was not influenced by the way in which the tumours were classified: benign, borderline and invasive malignant ovarian tumours; non-invasive (benign, borderline) and invasive; or benign and malignant (borderline and invasive malignant). However, the inter-observer agreement was reduced to good (Cohen’s Kappa 0.64-0.80) when the operators made probable diagnoses. In this subgroup of cases, agreement between the operators was best when ovarian tumours were dichotomised into non-invasive (benign and borderline) and invasive malignant lesions, as opposed to traditional diagnosis of benign and malignant (borderline and invasive cancers) lesions. This result indicates a tendency of the ultrasound operators to confuse benign tumours with borderline ovarian tumours.

In conclusion, our study has shown that the diagnostic performance with regard to discrimination between benign, borderline and malignant tumours when using pattern recognition depends on the confidence with which the diagnosis is suggested. Borderline ovarian tumours are the main source of diagnostic uncertainty. Our findings suggest that the level of confidence with which the diagnosis is made should be included into the ultrasound report. In cases of certain diagnosis, it would be safe to plan further management on the basis of the ultrasound diagnosis. In the case of a probable diagnosis, the morphological description of the tumour should be accompanied by a list of possible histological diagnoses and it might be wise to consider additional diagnostic methods to improve the diagnostic accuracy. In cases with uncertain diagnosis the ultrasound report should include a clear statement that ultrasound examination was not informative enough to establish the likely histological diagnosis.
Chapter 20

Study 6

Effect of the quality of gynaecological ultrasonography on management of patients with suspected ovarian cancer: a randomised controlled trial

This study shows that the quality of gynaecological ultrasonography diagnosis has a significant effect on the choices made by gynaecological oncologists in the management of patients with suspected ovarian cancer. The number of major staging surgical procedures for presumed ovarian malignancy was significantly lower after a level III (expert) ultrasonography compared with after a (routine) scan done by a non-specialist. This finding is likely to be the consequence of the greatly increased proportion of patients in whom a conclusive diagnosis of the nature of the adnexal tumour was possible from level III ultrasonography compared with level II ultrasonography. It could be argued that level II operators used a different ultrasound machine, which may be regarded by some ultrasound operators as being inferior to the one used by the level III operators in our study. Furthermore, level II operators could also have been at a different time pressure compared to level III operators. However, all the level II operators in our study were happy with the quality of the ultrasound machine they used and they always took all the time they needed to complete the examination.

The total number of surgical procedures done in the two groups was similar. However, the number of minimally invasive procedures in the Level III group was higher than in the Level II group, which contributed to the significant decrease in the length of hospital stay in the level III group.

A common problem in the management of patients with adnexal masses in routine clinical practice is the inability of the ultrasonography operator to reach a confident preoperative differentiation between benign and malignant masses. As a
result, benign cysts in many patients are treated as potentially malignant and these patients are routinely referred to regional cancer centres for further management. This problem is clearly highlighted in this study, where only 18 (12%) of the 150 of patients recruited were eventually diagnosed with ovarian cancer.

The main difference between the two groups of this study was the significantly higher proportion of non-diagnostic scans in the level II group. A likely histological diagnosis was provided to the clinicians in 76 (99%) of 77 patients in the level III group compared with 38 (52%) of 73 patients in the level II group. This finding can be largely explained by the differing approaches to obtaining a diagnosis from ultrasonography between non-specialist ultrasonographers and experts. Ultrasonographers tend to provide a summary of morphological characteristics of adnexal tumours on ultrasonography, but infrequently attempt to differentiate between benign and malignant ovarian tumours or attempt to predict a histological diagnosis. The interpretation of the ultrasonography findings is, thus, left to the clinicians with little or no experience in ultrasonography diagnosis, and who feel compelled to undertake surgery to ensure that patients with ovarian cancer are not managed incorrectly. The ultrasonography equipment used for each group was of a similar high quality and there is no reason to believe that the performance of ultrasonography machines contributed to the difference in group. This difference is more likely to be a result of discrepancies in the assessment of adnexal tumours and the reporting of ultrasonography findings between level II and level III operators.

Our findings suggest that another advantage of a level III scan is a better overall ability to diagnose benign adnexal pathology. This ability decreases the number of patients who are treated as potentially having ovarian cancer and aids the use of more conservative management options. Even in the context of care within the regional cancer centre, where the prevalence of malignancy is likely to be much higher compared
with low-risk units, the increased number of patients with a conclusive ultrasound

diagnosis of benign adnexal pathology was accompanied by a significant decrease in the

number of open staging surgical procedures.

The effect of expert scanning may be even higher if used in the primary

assessment of ovarian pathology. Increased confidence in the diagnosis of benign

ovarian lesions is likely to decrease the need for additional diagnostic tests, such as MRI

or serum CA-125 concentration, and also decrease the number of referrals to regional

cancer centres. A large proportion of benign ovarian cysts are detected incidentally

during opportunistic ultrasonography screening. Many of these patients are offered

surgical treatment because of uncertainty about the nature of the ovarian lesion, which

could be avoided by an improved ultrasonography diagnosis.

The accuracy of level III ultrasonography in this study was similar to the

findings reported from other expert centres (Valentin, 1999; Timmerman et al., 1999b).

One patient had a false negative diagnosis of squamous cell carcinoma, within a mature

cystic teratoma. This carcinoma is a rare tumour, which is difficult to diagnose. Expert

ultrasonography operators who use the “pattern recognition” method, do not normally

miss epithelial ovarian cancers. However, the diagnosis of borderline ovarian tumours is

more difficult and false negative findings are more common. This problem, however, is

offset by the high specificity achieved by expert ultrasonography operators, which helps

to avoid unnecessary interventions in patients with benign adnexal pathology.

The use of other imaging modalities to clarify the nature of the ovarian tumour

was no different between the two groups of our study, despite a large difference in the

number of conclusive ultrasonography diagnoses. This finding suggests clinicians

believe that other non-invasive tests contribute little to diagnostic accuracy, and they

tend to use surgery both as a diagnostic and a therapeutic procedure in ovarian tumours.

Such an approach is clearly not cost-effective and a powerful argument exists in favour
of improving the quality of ultrasonography assessment in routine gynaecological practice. However, the cost-benefit ratio might be affected by the greater cost of using expert ultrasonography operators compared with routine ultrasonographers.

In the UK, most ultrasonography assessments are done by ultrasonographers who are trained in the use of ultrasonography equipment for the assessment of patients with various medical conditions. The number of examiners who specialise in gynaecological ultrasonography is low. By contrast with other areas of clinical practice, such as obstetrics, the ultrasonography findings in gynaecology are often non-specific. To achieve a high diagnostic accuracy, ultrasonography imaging needs to be combined with clinical history and physical examination. This method has been recognised by the Royal College of Obstetricians and Gynaecologists who have integrated gynaecological ultrasonography training into core and advanced speciality training in gynaecology (Special Skills Training in Obstetrics and Gynaecology, 2002). This policy is likely to have a long term positive effect on the quality of diagnosis from gynaecological ultrasonography diagnosis in secondary and tertiary care centres. Until the number of highly trained gynaecological ultrasonographers is increased, effort should be made to provide expert ultrasonography scanning at least in tertiary referral units.
Chapter 21

Study 7

Value of preoperative ultrasound in the selection of women with adnexal masses for laparoscopic surgery

The results of our study have shown that preoperative ultrasound examination may be used to identify women with low risk of converting minimally invasive surgery for an ovarian tumour into a laparotomy. The most important prerequisite for performing safe laparoscopic surgery is an accurate diagnosis of a benign cyst (Parker, 1992). It is, therefore reassuring that there were no cases of invasive ovarian cancer in the study population.

One mucinous gastro-intestinal type borderline ovarian cyst was misdiagnosed as a benign tumour. The cyst appeared unilocular on the scan with a smooth internal wall and it did not contain solid foci or septations. At laparoscopy the cyst also appeared benign and the patient underwent a laparoscopic oophorectomy. She has been followed up with ultrasound scans, for the past two years, which showed no signs of tumour recurrence so far. A small number of gastro-intestinal type borderline ovarian tumours cannot be differentiated from benign cysts on ultrasound scan (Fruscella et al., 2005). Therefore it is likely that this type of borderline tumour will be occasionally operated on laparoscopically. However, the prognosis of gastro-intestinal type borderline ovarian tumours is very good, and there is no evidence that spillage of tumour content at laparoscopy has an adverse effect on long-term prognosis and survival rates (Ahmed and Lawton, 2005; Silverberg et al., 2004; Maneo et al., 2004). Therefore, the inability to rule out with absolute confidence a borderline gastro-intestinal tumour should not stop surgeons performing laparoscopies on women with presumed benign cysts. However, a careful audit of the outcome of these cases is
necessary in order to monitor for possible adverse outcomes that may be associated with spillage of borderline ovarian tumours.

Our study was in agreement with other reports (Valentin et al., 2001; Valentin, 2004), which examined the value of ultrasound pattern recognition for the diagnosis of ovarian cancer. As demonstrated in this study, pattern recognition was highly accurate in identifying benign lesions suitable for conservative surgery.

Berlanda et al. (2002) evaluated the accuracy and impact of a multiparameter, ultrasound-based triage on surgical management of adnexal masses. They divided tumours into low-risk, moderate-risk and high-risk of being malignant. However, three tumours that were diagnosed as low/moderate-risk turned out to be malignant on histological examination and they were classified as mismanaged, because they underwent laparoscopic surgery.

Another study by Guerriero and co-workers (2005) assessed whether ultrasonography with colour Doppler can identify and triage the patients with adnexal masses to the most appropriate surgical approach. Again, masses were classified based on the risk of malignancy. Adnexal masses with low risk of malignancy were operated on by conventional laparoscopy. No malignant tumour in this study was operated on by laparoscopic surgery.

However, our study did not attempt to differentiate adnexal tumours as being high or low risk of malignancy. All tumours suspected to be malignant were operated on by a gynaecological oncologist. Our study, aimed to predict the success of laparoscopic surgery by performing a preoperative ultrasound assessment, taking into account that these tumours must be benign in nature.

Approximately 80% of all presumed benign ovarian cysts were selected for a laparoscopy, which was successfully completed in nearly 95% of cases. This indicates that our criteria for performing laparoscopy may be used to identify those women in
whom the risk of conversion is low. It is equally important to emphasise that the proposed criteria are not too strict as almost four-fifths of women with benign cysts were offered minimally invasive surgery.

Our results also show that the sensitivity of ultrasound examination for the diagnosis of pelvic adhesions is not as high as we hoped. The PPV of ultrasound was not very good as in 2/6 (33%) of women with suspected severe adhesions the ultrasound diagnosis was not confirmed at laparotomy. Although only 3/113 (2.7%) women had severe adhesions that were not detected on the scan, adhesions were the main reason for failed laparoscopy and accounted for 50% of all failures. The presence of adhesions was particularly difficult to predict in obese patients and in those with large tumours. It is difficult to speculate why the accuracy of ultrasound was low in diagnosing pelvic adhesions in our study. The need for easier ultrasound signs that describe adhesions may be a solution. Nonetheless, the ultrasound examiner should maintain a high level of clinical suspicion and the assessment of pelvic anatomy’ mobility should be done systematically by a combination of gentle pressure with the vaginal probe and abdominal pressure with the examiners free hand as in a bimanual examination in order to assess movement of organs and tissues and their relation to each other.

Guerriero et al. (1997) in a previous study found that the adhesion of the ovary to the uterus (fixation), as evaluated by transvaginal ultrasonography, had a good PPV (81%) in diagnosing adhesions and this was more accurate than was blurring of the margins of the ovary (PPV 65%) and augmentation of the usual distance of the ovary from the probe which persisted with abdominal palpation (PPV 74%). In a recent study, Okaro and co-workers (2006) assessed ovarian mobility by applying pressure with the ultrasound probe. They reported a good correlation between ovarian mobility on transvaginal ultrasound and at laparoscopy (kappa 0.81). However, most of their
patients had normal ovaries, which are much easier to assess for mobility than are large pelvic tumours.

All the other unsuccessful laparoscopies failed purely because of technical difficulties, which may be encountered during laparoscopic surgery. Morbid obesity is not an absolute contraindication for laparoscopic surgery, but it is well known that it is often complicated by difficulties in gaining entry into the abdominal cavity. The other technical reason for failed laparoscopy in our study was an injury to the inferior epigastric artery. This is an unfortunate complication, which occasionally requires converting a laparoscopy into a laparotomy due to inability to contain bleeding from the injured vessel. Our conversion rate was not significantly different from the publication by Berlanda et al. (2002) and it was much lower than the 25% reported by Havrilesky (2003), who evaluated the clinical outcomes of laparoscopic management of adnexal masses that were thought to be benign before surgery. These differences are dependent on many factors including patient population, indication for surgery and experience of surgeon. It is difficult to compare our study with previous reports because we performed surgery only on women with conclusive diagnosis of benign ovarian tumours, whereas all other studies included women with significant proportion of malignant adnexal tumours.

It is also important to say that ultrasound was accurate in excluding severe pelvic endometriosis; hence there were no cases of severe endometriosis among those women who underwent intermediate-level laparoscopic surgery.

Laparoscopic ovarian cystectomy was successfully completed in just over 90% of the planned cases. Undiagnosed pelvic adhesions were the reason for conversion of laparoscopic cystectomy into laparoscopic oophorectomy in nearly a third of the cases. The remaining reasons for conversion were heavy bleeding from the cyst bed and the presence of a chronic adnexal torsion. However, ovarian tumour volume appeared to
have played an important role in dictating the success of laparoscopic ovarian
cystectomy as demonstrated in our results. Although it is difficult to establish a cut-off
tumour volume for laparoscopic ovarian cystectomy, we should bear in mind that the
risk of oophorectomy increases with the increasing size of the cyst. This is in agreement
with a previous study by Havrilesky et al. (2003), which also showed that the risk of
failed laparoscopic surgery increases with increasing tumour size.

There is no doubt that the success of laparoscopic surgery is mainly dependent
on the skill and expertise of the operating surgeon. In expert hands most operations can
be completed successfully using a laparoscopic approach. Most consultant
gynaecologists in the UK and worldwide perform intermediate-level laparoscopic
surgery. The results of this study may aid them when selecting women with benign
cysts for minimally invasive surgery.

In conclusion, our study has shown that a detailed preoperative transvaginal
ultrasound is helpful for assessing the feasibility of intermediate-level laparoscopic
surgery in women with benign adnexal lesions. A correct diagnosis of benign tumour is
of critical importance when deciding on the type of operation to remove an adnexal
cyst. However, the assessment of tumour consistency and mobility helps to identify
women in whom the risk of conversion to laparotomy is minimal and surgery may be
performed as a day case.
PART FOUR:

Conclusions and Further Research
Chapter 23: Conclusions and Further Research

This thesis has shown that RMI and OCS are fairly good methods for preoperative differentiation between invasive and non-invasive adnexal tumours. Both tests have proven to be very sensitive for the detection of invasive ovarian cancer. The use of the OCS is easy and has the potential benefit of reducing the number of unnecessary referrals to regional gynaecological cancer centres for planning further management, and it would eventually lead to reduction in both cost and anxiety. However, future work including a higher number of ovarian malignancies, will show whether the use of the OCS as an alternative to RMI, either during the initial scan or within tertiary referral centres, may be a more effective way of managing women with ultrasonographic diagnosis of ovarian tumours.

This work had demonstrated that the correct ultrasound diagnosis of BOT using pattern recognition can be achieved in two-thirds of cases with a very good specificity. Further work is required to identify morphological and biochemical features that may help to increase the accuracy of pre-operative detection of such tumours.

The thesis has confirmed that preoperative diagnosis of an adnexal mass made on the basis of a real-time ultrasound examination is more accurate than a diagnosis made on the basis of saved static ultrasound images. Evaluation of static images is associated with lower specificity. The dynamic aspect of the real-time scan gives the examiner supplementary information and makes it possible to scrutinise every part of the mass. The ‘rarity’ of experts in gynaecological ultrasonography would make the aspect of providing a second opinion by email, in the case of ‘difficult’ adnexal tumours, very appealing. It would be interesting in the future to assess whether the performance of these experts would be better if they were presented with representative
video clips or with the use of four-dimensional cine loops of volume instead of with static images.

Our dataset showed that ultrasound pattern recognition is an accurate and highly reproducible method for characterisation of adnexal tumours, and its accuracy is dependent on the confidence with which the diagnosis is suggested. The test, however, performs less well in cases of BOT, which are sometimes difficult to distinguish from benign tumours and are the main source of diagnostic uncertainty. We suggest that the structure of reporting ultrasound findings may need to change in order to classify ovarian tumours as being non-invasive or invasive, rather than making a diagnosis of benign or malignant tumour. It is also important that the level of confidence with which the diagnosis is made is also included into the ultrasound report. In cases of certain diagnosis, it would be safe to plan further management on the basis of the ultrasound diagnosis. In the case of a probable diagnosis, the morphological description of the tumour should be accompanied by a list of possible histological diagnoses and it might be wise to consider additional diagnostic methods to improve the diagnostic accuracy. In cases with uncertain diagnosis the ultrasound report should include a clear statement that ultrasound examination was not informative enough to establish the likely histological diagnosis.

The randomised controlled trial included in this thesis has confirmed the importance of expertise in gynaecological ultrasonography in view of the improved accuracy in differential diagnosis of ovarian tumours and the positive effect on their optimal management. Our findings suggest that a greater number of minimally invasive procedures and shorter hospital stays can be achieved in women with adnexal lesions suspicious of ovarian malignancy if they are scanned by experienced gynaecologists rather than ultrasonographers. There is an urgent need for a detailed cost–benefit analysis and quality of life assessment of this management strategy and for its
assessment in the primary-care setting. Independent validation should lead to the
provision of expert ultrasonography, at least in tertiary cancer centres. Additionally, in
countries like the UK, where few gynaecologists specialise in ultrasonography and most
scans are done by ultrasonographers, there are important implications for training and
workforce planning.

Finally, this thesis did show that detailed preoperative transvaginal ultrasound is
helpful for assessing the feasibility of intermediate-level laparoscopic surgery in women
with benign adnexal lesions. A correct diagnosis of benign tumour is of critical
importance when deciding on the type of operation to remove an adnexal cyst.
However, the assessment of tumour consistency and mobility helps to identify women
in whom the risk of conversion to laparotomy is minimal and surgery may be performed
as a day case. Future work should concentrate on assessing the role of gynaecological
ultrasonography in the estimation of the severity of pelvic endometriosis. The
identification of patients with severe endometriosis will provide these women with an
opportunity to be managed by experts in laparoscopic surgery in tertiary endometriosis
centres.
Bibliography


- Aslam N, Ong C, Woelfer B, Nicolaides K, Jurkovic D 2000b Serum CA125 at 11-14 weeks of gestation in women with morphologically normal ovaries. BJOG 107:689-90

- Athey PA, Cooper NB 1985 Sonographic features of paraovarian cysts. AJR Am J Roentgenol 144:83-6


• Buamah P 2000 Benign conditions associated with raised serum CA125 concentration. J Surg Oncol 75:264-5


202


• Gotlieb WH, Chetrit A, Menczer J, Hirsh-Yechezkel G, Lubin F, Friedman E, Modan B, Ben-Baruch G 2005 Demographic and genetic characteristics of patients with borderline ovarian tumors as compared to early stage invasive ovarian cancer. Gynecol Oncol 97:780-3


• Granberg S, Wikland M, Jansson I 1989 Macroscopic characterisation of ovarian tumors and the relation to the histological diagnosis: Criteria to be used for ultrasound evaluation. Gynecol Oncol 35:139–44
• Granberg S, Norstrom A, Wikland M 1990 Tumors in the lower pelvis as imaged by vaginal sonography. Gynecol Oncol 37:224-9


• Guerriero S, Ajossa S, Garau N, Piras B, Paoletti AM, Melis GB 2005 Ultrasonography and color Doppler-based triage for adnexal masses to provide the most appropriate surgical approach. Am J Obstet Gynecol 192:401-6


• Hacker NF, Berek JS, Lagasse LD 1983 Primary cytoreductive surgery for epithelial ovarian cancer. Obstet Gynecol 61:413-20


204


• Hart WR 2005 Mucinous tumors of the ovary: a review. Int J Gynecol Pathol 24:4-25


• Hillaby K, Aslam N, Salim R, Lawrence A, Raju KS, Jurkovic D 2004 The
value of detection of normal ovarian tissue (the ‘ovarian crescent sign’) in the

• Hoffer FA, Kozakewich H, Colodny A, Goldstein DP 1988 Peritoneal inclusion

Shiozawa I, Ueda N, Konishi I 2003 Toward understanding the natural history
of ovarian carcinoma development: a clinicopathological approach. Gynecol
Oncol 88:309-17

• Howard FM 1995 Surgical management of benign cystic teratoma: Laparoscopy
vs. laparotomy. J Reprod Med 40:495–9

• Huus M, Lafay-Pillet MC, Lecuru F, Ruscillo MM, Chevalier JM, Vilde F,
Taurelle R 1996 Granulomatous peritonitis after laparoscopic surgery of an
ovarian dermoid cyst. Diagnosis, management, prevention, a case report. J

• International Federation of Gynecology and Obstetrics: Classification and
staging of malignant tumors in the female pelvis 1971 Acta Obstet Gynecol
Scand 50:1-7

Worldwide (2002 estimates)

• ISD Scotland Online 2007 Cancer Incidence, Mortality and Survival data

• Jacobs IJ, Fay TN, Stabile I, Bridges JE, Oram DH, Grudzinskas JG 1988 The
distribution of CA 125 in the reproductive tract of pregnant and non-pregnant
women Br J Obstet Gynaecol 95:1190-4

• Jain KA 1994 Prospective evaluation of adnexal masses with endovaginal grey-scale and duplex and color Doppler US: correlation with pathologic findings. Radiology 191: 63-7

• Jain KA 2000 Imaging of peritoneal inclusion cysts. AJR Am J Roentgen 174:1559-63

• Jaeschke R, Guyatt GH, Sackett DL (For the Evidence-Based Medicine Working Group) 1994 User's guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? JAMA 271:703-7


207


- Levine D 1994 What is the significance of the incidental discovery of a unilocular ovarian cyst in a postmenopausal woman (either with or without a family history of ovarian cancer) during a pelvic sonographic examination to exclude ovarian cancer? AJR 163, 215–216


• Newey VR 1997 Classical versus artificial neural network analysis (Opinion) Ultrasound Obstet Gynecol 10:5-8

• Northern Ireland Cancer Registry 2007 Cancer Registrations in Northern Ireland, 2004


• Okaro E, Condous G, Khalid A, Timmerman D, Ameye L, Huffel SV, Bourne T 2006 The use of ultrasound-based ‘soft markers’ for the prediction of pelvic pathology in women with chronic pelvic pain—can we reduce the need for laparoscopy? Br J Obstet Gynaecol 113:251-6


• Patel MD, Feldstein VA, Chen DC Lipson SD, Filly RA 1999 Endometriomas: diagnostic performance of US. Radiology 210: 739-45


• RCOG special skills training modules 2002 Ultrasound imaging in the management of gynaecological conditions. RCOG press

• RCOG Greentop guidelines 2003 Ovarian cysts in postmenopausal women. RCOG. Guideline No. 34


• Roman LD 1998 Small cystic pelvic masses in older women: is surgical removal necessary? Gynecol Oncol 69:1-2
• Russell P 1979 The pathological assessment of ovarian neoplasms. I: Introduction to the common 'epithelial' tumours and analysis of benign 'epithelial' tumours. Pathology 11:5-26
• Sample WF, Lippe BM, Gyepes MT 1977 Gray-scale ultrasonography of the normal female pelvis. Radiology 125:477-83


• Serov SF, Scully RE, Sobin LH 1973 Histological Typing of Ovarian Tumours. (WHO International Histological Classification of Tumours No. 9). World Health Organization: Geneva, Switzerland


• Shih IeM, Kurman RJ 2004 Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. Am J Pathol 164:1511-8
• Shiozawa I, Ueda N, Konishi I 2003 Toward understanding the natural history of ovarian carcinoma development: a clinicopathological approach. Gynecol Oncol 88:309-17


• Troiano RN, McCarthy S 1994 Magnetic resonance imaging evaluation of adnexal masses. Semin Ultrasound CT MR 15:38–48


• Valentin L 1999 Prospective cross-validation of Doppler ultrasound examination and gray-scale ultrasound imaging for discrimination of benign and malignant pelvic masses. Ultrasound Obstet Gynecol 14: 273-83

• Valentin L 2000 Comparison of Lerner score, Doppler ultrasound examination, and their combination for discrimination between benign and malignant adnexal masses. Ultrasound Obstet Gynecol 15:143-7


• Valentin L, Ameye L, Jurkovic D, Metzger U, Lécuru F, Van Huffel S, Timmerman D 2006b Which extraterine 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 Obstet Gynecol 27: 438–44


- Welsh Cancer Intelligence and Surveillance Unit 2007
- Wu CC, Lee CN, Chen TM, Shyu MK, Hsieh CY, Chen HY, Hsieh FJ 1994 Incremental angiogenesis assessed by color Doppler ultrasound in the tumorigenesis of ovarian neoplasms. Cancer 73;1251-6
Publications

The results of this thesis have generated the following publications:


The following studies have been accepted for publication in the journal of Ultrasound in Obstetrics and Gynecology:
