CORE
Metadata, citation and similar papers at core.ac.uk
Provided by University of Birmingham Research Archive, E-theses Repository

EVALUATION OF AMBULATORY DIAGNOSIS OF ABNORMAL UTERINE BLEEDING

by

SHAGAF HAJ BAKOUR

A thesis submitted to the Faculty of Medicine and Dentistry

Of the University of Birmingham

For the degree of

DOCTOR OF MEDICINE

Department of Obstetrics and Gynaecology
Faculty of Medicine and Dentistry
The University of Birmingham
September 2002

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

In the ambulatory assessment of women with abnormal uterine bleeding, the main aim is to reach a diagnosis and thereby allow appropriate treatment. Excluding endometrial pathology, particularly carcinoma and hyperplasia, is of paramount importance. However there is no consensus as to which set of investigations should be used (hysteroscopy, ultrasonography and endometrial sampling). There is a lack of good quality research evidence on the accuracy of the various diagnostic tests in predicting endometrial lesions. This dearth of relevant evidence prompted the research presented in this thesis to address and answer the following research questions:

- What is the accuracy of outpatient miniature hysteroscopy / ultrasonography in the identification of premalignant and malignant endometrial lesions?
- What is the relative significance of hysteroscopic and ultrasonographic evidence of endometrial atrophy in relation to an insufficient sample on outpatient endometrial biopsy?
- What is the risk of premalignant and malignant pathology among endometrial polyps? What is the significance of various risk factors associated with endometrial polyps?
- What is the feasibility of multivariable analysis to evaluate combinations of diagnostic tests in the diagnosis of endometrial disease?

Findings and Conclusions

Positive hhysteroscopy is accurate in ruling in endometrial cancer and hyperplasia (the LR was 51.1 (95% CI 7.9-326.9). Using endometrial thickness $\geq$4mm at

ultrasound scan, ultrasound is accurate in ruling out endometrial cancer and hyperplasia (the LR was 0.14 (95%CI 0.02-0.64).

Insufficient sample on endometrial biopsy was more common among cases with hysteroscopic finding of endometrial atrophy (adjusted OR= 4.79, 95% CI 1.05-21.91, p=0.04) and less common among cases with sonographic endometrial thickness above 5mm (adjusted OR=0.19, 95% CI 0.07-0.53, p=0.001). Therefore insufficient sample may be considered a substitute to absence of pathology provided the hysteroscopic and sonographic endometrial assessment is consistent with endometrial atrophy.

Hyperplasia was more frequent in endometrial specimens with polyps than in those without (9.7% vs 4.8%, OR=2.1, 95% CI 0.85-5.2), but the rate of carcinoma in the two groups was not statistically different (4.8% vs 3.2%, OR=1.5, 95% CI 0.46-5.0). Tamoxifen treatment was associated with endometrial polyps (adjusted OR= 8.16, 95% CI 2.01-33.2, p=0.003) but hormone replacement therapy was not (adjusted OR=1.35, 95% CI 0.65 – 2.81, p=0.42).

The true clinical value of a test lies in the added information over and above what was already known from the history and examination. It is feasible to develop a stepwise multivariable analytic approach to explore the added value of tests (hysteroscopy or ultrasonography) over and above history when predicting endometrial hyperplasia or cancer. This analytic strategy needs to be applied in larger datasets to draw clinical conclusions.

DEDICATION

I wish to dedicate this thesis to my dear parents and my dear husband Dham

ACKNOWLEDGEMENTS

This work was undertaken at the One-stop Hysteroscopy clinic at the Birmingham Women's Hospital between November 1996 and December 1997. I would like to acknowledge and thank those who have assisted me during that time to make this work possible.

I am particularly indebted to my supervisor, Mr Khalid S. Khan, for his critical guidance, and continuous support he has given me throughout my career. I also wish to thank Mr Janesh Gupta, senior lecturer, Minimal Access Surgery, for his help and advice, and Mr David Pollard for dedicating the time and effort to review the English language of the thesis. In addition, I would like to thank the patients attending the One-stop Hysteroscopy clinic whose data were used in this thesis.

And the last but not the least I would like to thank the clerical, clinical and nursing staff at the One-stop Hysteroscopy clinic for their effort with the work involved in this thesis.

TABLE OF CONTENTS

		PAGE
EXECUTIVE ABSTRACT		**1**
Background		1
Objectives		1
Methods		2
Findings		3
Conclusions		4
CHAPTER 1:	**INTRODUCTION**	**5**
1.1	Definition	5
1.2	Prevalence	5
1.3	Prevalence of endometrial hyperplasia and cancer	6
1.4	Evaluating women with abnormal uterine bleeding	6
1.5	Hysteroscopy	7
1.6	Ultrasonography	8
1.7	Outpatient endometrial biopsy	8
1.8	Outpatient setting	9
1.9	Literature review	9
1.10	Endometrial polyps	13
1.11	Insufficient sample	13
1.12	Questions addressed in this thesis	14

CHAPTER 2: THE DIAGNOSTIC ACCURACY OF OUTPATIENT MINATURE HYSTEROSCOPY IN PREDICTING PREMALIGNANT AND MALIGNANT ENDOMETRIAL LESIONS **15**

2.1 Introduction 15

2.2 Ethical approval 17

2.3 Methods 17

2.4 Results 20

2.5 Discussion 25

2.5.1 Principal findings of the study 25

2.5.2 Appraisal of quality of the study 25

2.5.3 Relation to other studies considering differences in quality and results 28

2.5.4 Practical implications for clinicians 29

2.5.5 Implications on research 29

CHAPTER 3: THE DIAGNOSTIC ACCURACY OF ULTRASOUND SCAN IN PREDICTING ENDOMETRIAL HYPERPLASIA AND CANCER IN POSTMENOPAUSAL BLEEDING **31**

3.1.1 Introduction 31

3.1.2 The effect of hormone replacement therapy on the endometrium 32

3.2 Methods 34

3.3 Results 36

3.4 Discussion 40

3.4.1 Principal findings of the study 40

3.4.2 Appraisal of quality of the study 40

3.4.3 Relation to other studies considering differences in quality and results 42

3.4.4 Practical implications for clinicians 45

3.4.5 Implications on research 45

CHAPTER 4: CONTROLLED ANALYSIS OF FACTORS ASSOCIATED WITH INSUFFICIENT SAMPLE ON OUTPATIENT ENDOMETRIAL BIOPSY **46**

4.1 Introduction 46

4.2 Methods 47

4.3 Results 48

4.4 Discussion 52

4.4.1 Principal findings of the study 52

4.4.2 Appraisal of quality of the study 52

4.4.3 Relation to other studies considering differences in quality and results 54

4.4.4 Practical implications for clinicians 54

4.4.5 Implications on research 54

CHAPTER 5: THE RISK OF PREMALIGNANT AND MALIGNANT PATHOLOGY IN ENDOMETRIAL POLYPS **56**

5.1 Introduction 56

5.2 Methods 57

5.3 Results 58

5.4 Discussion 62

5.4.1 Principal findings of the study 62

5.4.2 Appraisal of quality of the study 62

5.4.3 Relation to other studies considering differences in quality and results 63

5.4.4 Practical implications for clinicians 64

5.4.5 Implications on research 64

CHAPTER 6: RISK FACTORS ASSOCIATED WITH ENDOMETRIAL POLYPS IN ABNORMAL UTERINE BLEEDING **65**

6.1 Introduction 65

6.2 Methods 65

6.3 Results 67

6.4 Discussion 71

6.4.1 Principal findings of the study 71

6.4.2 Appraisal of quality of the study 71

6.4.3 Relation to other studies considering differences in quality and results 73

6.4.4 Practical implications for clinicians 73

6.4.5 Implications on research 74

CHAPTER 7: EVALUATION OF OUTPATIENT HYSTEROSCOPY AND ULTRASONOGRAPHY IN THE DIAGNOSISOF ENDOMETRIAL DISEASE 75

7.1 Introduction 75

7.2 Methods 76

7.3 Results 79

7.4 Discussion 84

7.4.1 Principal findings of the study 84

7.4.2 Appraisal of quality of the study 84

7.4.3 Relation to other studies considering differences in quality and results 87

7.4.4 Practical implications for clinicians 87

7.4.5 Implications on research 87

CHAPTER 8: DISCUSSION AND CONCLUSIONS 88

8.1 The principle findings of this thesis 88

8.2 Appraisal of quality of the study 91

8.3 Practical implications for clinicians 95

8.4 Implications on research 98

APPENDIX 1: Letters from the Chief of Research and Development Committee and from the Clinical Governance Manager regarding the ethical approval 99

APPENDIX 2: Clarification about author's contribution to chapter 3 101

REFERENCES **102**

PUBLICATIONS FROM THE THESIS **118**

LIST OF TABLES

Chapter 1

		Page
Table 1.1	Conclusions of reviews on accuracy of tests for menstrual disorders	10
Table 1.2	Quality assessment of reviews on accuracy of tests for menstrual disorders. Aadopted from the QUOROM checklist	12

Chapter 2

Page

Table 2.1	Diagnostic value of hysteroscopy in prediction of endometrial premalignant/malignant lesion.	22
Table 2.2	Diagnostic value of hysteroscopy in prediction of endometrial premalignant/malignant lesion in postmenopausal and premenopausal women	24
Table 2.3	Application of STARD checklist to the study reported in Chapter 2	27

Chapter 3

Page

Table 3.1	Diagnostic value of ultrasound scan in diagnosis of endometrial cancer and hyperplasia	39
Table 3.2	Diagnostic value of ultrasound scan in diagnosis of endometrial cancer or hyperplasia	40
Table 3.3	Application of STARD checklist to the study reported in Chapter 3	42

Chapter 4

Page

Table 4.1	Characteristics of women attending outpatient One-stop hysteroscopy clinic	50
Table 4.2	Logistic regression analysis to determine the relationship between various factors associated with insufficient specimen on outpatient endometrial biopsy	52
Table 4.3	Application of the Standard of Quality for Reporting Logistic Regression Models	54

Chapter 5

Page

Table 5.1	Endometrial hyperplasia and cancer in women with endometrial polyps and those without	59
Table 5.2	Histology of endometrial polyps and estrogenic exposure to hormone replacement therapy or tamoxifen	61

Chapter 6

Page

Table 6.1	Comparison of women with and without polyps diagnosed in One-stop clinic	68
Table 6.2	Logistic regression analysis to determine the significance of factors associated with endometrial polyps	70
Table 6.3	Application of the Standard of Quality for Reporting Logistic Regression Models	72

Chapter 7

Page

Table 7.1	Results of univariable and stepwise multivariable analyses to assess the ability of patient history, hysteroscopy and ultrasonography to predict the presence of endometrial hyperplasia or cancer in 248 women with abnormal uterine bleeding.	81
Table 7.2	Application of the Standard of Quality for Reporting Logistic Regression Models	86

Chapter 8

Page

Table 8.1	Summary of the thesis results	90
Table 8.2:	Application of STARD checklist to the study reported in this thesis	92
Table 8.3:	Classification of level of evidence for primary research on diagnostic accuracy	94

LIST OF FIGURES AND CHARTS

Chapter 7

		Page
Figure 7.1	Improvement in diagnostic accuracy achieved by addition of tests to the information provided by patient history. Area under receiver operating characteristic (ROC) curves of the diagnostic models based on patient history, patient history plus endometrial thickness measured by ultrasound scan and patient history plus ultrasound scan and hysteroscopy	82
Figure 7.2	Improvement in diagnostic accuracy achieved by addition of tests to the information provided by patient history. Area under receiver operating characteristic (ROC) curves of the diagnostic models based on patient history, patient history plus hysteroscopy and patient history plus hysteroscopy and endometrial thickness measured by ultrasound scan	83

Chapter 8

Flow Chart	Management of women with abnormal uterine bleeding	97

LIST OF ABBREVIATIONS

CI	Confidence Intervals
dOR	diagnostic odds ratio
ET	Endometrial thickness
GMAU	Gynaecology Minimal Access Unit (GMAU), Ninewells Hospital, Dundee
HRT	Hormone replacement therapy
LRs	Likelihood Ratios
MAST centre	Birmingham Minimal Access and Surgical Training
mm	Millimetre
OR	Odds Ratio
QUOROM	Quality of Reporting of Meta-analysis
PMB	Post menopausal bleeding
ROC	Receiver operating characteristic curve
STARD group	Standards for Reporting of Diagnostic Accuracy

EXECUTIVE ABSTRACT

Background

In ambulatory assessment of women with abnormal uterine bleeding, the main aim is to exclude endometrial pathology, particularly carcinoma and hyperplasia. However there is no consensus as to which set of investigations should be used. While some clinicians resort to ultrasound imaging techniques as well as tissue sampling, others advocate endoscopic examination. There is a lack of good quality research evidence on the accuracy of the various diagnostic tests in predicting endometrial lesions. This dearth of relevant evidence prompted the research presented in this thesis.

Objectives

This thesis addresses the following research questions:

1. What is the accuracy of outpatient miniature hysteroscopy in the identification of premalignant and malignant endometrial lesions?
2. What is the accuracy of ultrasound scan in the identification of premalignant and malignant endometrial lesions?
3. What is the relative significance of hysteroscopic and ultrasonographic evidence of endometrial atrophy in relation to insufficient sample on outpatient endometrial biopsy?
4. What is the risk of premalignant and malignant pathology among endometrial polyps?
5. What is the significance of various risk factors associated with endometrial polyps?
6. What is the feasibility of multivariable analysis to evaluate combinations of diagnostic tests in the diagnosis of endometrial disease?

Methods

A prospective cohort study was performed on 248 patients evaluated for abnormal uterine bleeding in the One-stop clinic at Birmingham Women's Hospital from November 1996 to December 1997. Outpatient hysteroscopy, ultrasonography and endometrial sampling were performed. Histological diagnosis was used as a "gold standard" for cancer and hyperplasia. The histological nature of the polyp was confirmed in all patients by performing inpatient polypectomy.

For each question, the methodology was as follows:

1. For the diagnostic value of hysteroscopy, a test accuracy study was carried out where likelihood ratios (LRs) and post-test probabilities were computed along with confidence intervals (CI).
2. For the diagnostic value of ultrasonography, a test accuracy study was carried out where LRs and post-test probabilities were computed along with confidence intervals.
3. Multivariable logistic regression modelling was used to evaluate the independent association of hysteroscopic findings and sonographic endometrial thickness with insufficient outpatient endometrial sampling controlling for the effects of age and menopausal status.
4. To determine the magnitude of malignant potential among polyps, the odds for pathological findings in polyps (cases) were statistically compared with the odds amongst non-polypoidal specimens (controls).
5. Multivariable logistic regression modelling was used to evaluate the effects of age, parity, menopausal status, hormone replacement therapy and tamoxifen treatment on the occurrence of polyps.
6. Multivariable logistic regression models were built in a stepwise fashion, to test whether ultrasonography and/or hysteroscopy increased the prediction of serious endometrial pathology above that predicted from clinical history alone. The diagnostic accuracy of the

models was determined by the area under the receiver operating characteristic curve.

Findings

1. For normal findings at hysteroscopy the LR was 0.7 (95% CI 0.5-0.9). With definitely abnormal findings at hysteroscopy, the LR was 51.1 (95% CI 7.9-326.9) for predicting premalignant and malignant endometrial lesions.
2. For endometrial thickness ≤4mm at ultrasound scan, the LR was 0.14 (95%CI 0.02-0.64). It was 0.94 (95%CI 0.37-1.99) for endometrial thickness 4.1-9.0mm, and 3.3 (95%CI 1.73-5.73) for endometrial thickness >9.0mm for predicting premalignant and malignant endometrial lesions.
3. Insufficient sample on endometrial biopsy was more common among cases with hysteroscopic finding of endometrial atrophy (adjusted OR= 4.79, 95% CI 1.05-21.91, p=0.04) and less common among cases with sonographic endometrial thickness above 5mm (adjusted OR= 0.19, 95% CI 0.07-0.53, p=0.001).
4. Hyperplasia was more frequent in endometrial specimens with polyps than in those without (9.7% vs 4.8%, OR=2.1, 95% CI 0.85-5.2), but the rate of carcinoma in the two groups was not statistically different (4.8% vs 3.2%, OR=1.5, 95% CI 0.46-5.0).
5. Tamoxifen treatment was associated with endometrial polyps (adjusted OR= 11.21, 95% CI 2.70 – 46.46, p=0.0009) but hormone replacement therapy was not (adjusted OR= 1.48, 95% CI 0.68 - 3.20, p=0.32).
6. The area under the receiver operating characteristic curve for the model including historical features alone was 0.78. When hysteroscopy and ultrasonography were added to the model, the ROC area increased to 0.81 (p=0.008 for improvement) and 0.82 (p=0.02 for improvement) respectively.

Conclusions

1. Hysteroscopy is accurate in ruling in the diagnosis of endometrial cancer and hyperplasia, but not in ruling it out.
2. Ultrasound is accurate in ruling out the diagnosis of endometrial cancer and hyperplasia, but not in ruling it in.
3. Insufficient sample on outpatient endometrial biopsy can be anticipated when hysteroscopic and sonographic endometrial assessment is consistent with endometrial atrophy and in these cases it is not associated with pathology.
4. The rate of hyperplasia in endometrial polyps is approximately double that in the non-polypoidal endometrium while the cancer rate is no different.
5. Tamoxifen treatment is associated with endometrial polyps but hormone replacement therapy is not.
6. The true clinical value of a test lies in the added information over and above what was already known from the history and examination. It is feasible to develop a stepwise multivariable analytic approach to explore the added value of tests (hysteroscopy or ultrasonography) over and above history when predicting endometrial hyperplasia or cancer. This analytic strategy needs to be applied in larger datasets to draw clinical conclusions.

CHAPTER 1

INTRODUCTION

1.1 Definition

Abnormal uterine bleeding refers to a range of abnormal premenopausal or postmenopausal uterine bleeding symptoms and in women of reproductive age is defined as bleeding at abnormal or unexpected times or by excessive flow (>80 mls) at the time of an expected menses. Any bleeding should be considered abnormal in post-menopausal women except for those with predictable withdrawal bleeding taking hormone replacement therapy (Goodman 2000). In this thesis, the term abnormal uterine bleeding will include the whole range of premenopausal menstrual disorders, postmenopausal bleeding and unscheduled bleeding on hormone replacement therapy (HRT).

1.2 Prevalence

Abnormal uterine bleeding is one of the most common gynaecologic problems that health care providers will face, accounting for approximately 15-20% of office visits and 25% of gynaecologic operations (Goodman 2000). It has an impact on many women's quality of life. Once referred to a gynaecologist, surgical intervention is highly likely. Menstrual disorders occur secondary to a wide variety of functional and structural abnormalities. One in five women in the UK will have a hysterectomy before the age of 60 years. Each year around £7 million is spent on primary care prescriptions to treat menorrhagia, (Royal College of Obstetrics and Gynaecologists guidelines 1998). There may be cost savings with a reduction in unnecessary tests or with the avoidance of inappropriate surgical intervention if the investigation process can be rationalised.

1.3 Prevalence of endometrial hyperplasia and cancer

Diagnostic evaluation of abnormal uterine bleeding is carried out mainly to exclude endometrial carcinoma, although a malignancy is found in <10% of the cases (Lidor 1986). Only up to 10% to 20% of cases of endometrial cancer occur before menopause, but most of these women are aged between 40 and 50 years (Gallop 1984; Jeffery *et al.,* 1987). In the United Kingdom the predicted frequency of endometrial cancer in women aged <40 years is <12 cases per 100,000 women (Gallop 1984; Jeffery *et al*., 1987; Rose 1996), whereas in women aged <20 years it is one case per 208,000 women (Fraser 1972).

Endometrial hyperplasia is considered to be a precursor for endometrial cancer and in the absence of therapy, progression to endometrial cancer occurs in up to 3% to 23% of cases of complex hyperplasia over a 13-year period (Terakawa *et al*., 1997). Those cases of hyperplasia with atypia without treatment are more likely to progress to cancer (up to 75%) than other types of hyperplasia, and a proportion (estimated at approximately 25%) will have undetected carcinoma of the endometrium at the time of the initial biopsy (Farquhar *et al.,* 1999; Hunter 2001; Kurman *et al.,* 1985).

1.4 Evaluating women with abnormal uterine bleeding

Establishing diagnosis is crucial to clinical management as it allows one to choose appropriate treatment. Therefore evidence-based diagnosis is part of evidence-based practice; and in establishing accurate diagnosis, the true clinical value of a test lies in the added information over and above what is already known from the history and examination.

Conducting investigations and establishing a diagnosis should be accomplished without delay particularly in women over the age of 40 years preferably in an out patient setting. This is because

there is a need to exclude endometrial cancer in this age group. Providing this service in out patient setting is both efficient and effective in that the diagnosis and the treatment (in some cases) can be established in one-stop consultation. This would save patients going to theatre to have the same procedures performed but under general anaesthesia and would further result in time saving of staff and patient alike, as well as less domestic and occupational disruption.

The key to accurate, efficient diagnosis of disorders responsible for abnormal uterine bleeding depends on a systematic consideration of all the possible causes. Appropriate workup should be guided by age-related differential diagnoses for abnormal bleeding. Careful history and physical and pelvic examinations provide the framework for evaluation. Suspicion of reproductive tract malignancies is heightened in patients >40 years old, those with a history of oligoovulation or anovulation suggestive of long-term unopposed estrogenic exposure, those who are obese, and those who do not respond to first-line medical management (Long 1996). Modern diagnostic tools can quickly focus the evaluation and allow timely intervention. The availability of these non-invasive and minimally invasive diagnostic tools and minimally invasive surgical treatment has revolutionized the management of abnormal uterine bleeding. Advancements in technology have allowed considerable improvements in the resolution of many diagnostic tools. The adjunctive diagnostic techniques available include hysteroscopy, transvaginal ultrasonography, endometrial biopsy, and dilatation and curettage.

1.5 Hysteroscopy

Hysteroscopy is an important means of visualising intrauterine pathologic disorders in patients with various gynaecologic complaints, including menstrual disorders. Successful visualisation of the endometrial cavity in office hysteroscopy depends on the appropriate selection of the patient, the absence of contraindications, and adequate instrumentation and technique. Although clinicians were

limited to blind endometrial curettage or less effective imaging techniques in the past, we are now able to thoroughly evaluate the intrauterine environment with direct visualisation of the uterine cavity using the hysteroscope as an invaluable adjunct to blind endometrial sampling. Prior reports have revealed that in 60% of hysterectomy specimens obtained immediately after curettage the source of excessive bleeding was frequently not diagnosable on the curettage specimen (Stock 1975; Word 1958). Consequently, diagnostic hysteroscopy has become increasingly popular over the past 20 years. In 1990 de Jong (1990) showed that office hysteroscopy under local anaesthesia is a reliable method for assessing the uterine cavity and it causes little discomfort (Towbin 1996). However, its diagnostic accuracy has not yet been established with methodologically sound research.

1.6 Ultrasonography

Ultrasonographic studies (particularly transvaginal), have played an increasing role in the evaluation of patients with menstrual disorders / abnormal uterine bleeding over the past decade, especially for cases of intrauterine space-occupying lesions, including endometrial polyps, and submucosal myomas. The accuracy of this technology has also been poorly evaluated in research.

1.7 Outpatient endometrial biopsy

Endometrial biopsy has been claimed to be a sensitive and relatively inexpensive test for identifying endometrial hyperplasia and carcinoma, with a reported sensitivity for the detection of endometrial carcinoma of approximately 85% to 95% (Lidor 1986). However, endometrial biopsy is deemed a

poor test for diagnosing other benign endometrial abnormalities, such as polyps and submucosal leiomyomas (O'Connell *et al.,* 1998).

1.8 Outpatient setting

In the ambulatory assessment and management of women with abnormal uterine bleeding, there is no consensus as to which set of investigations should be implemented. While some clinicians would use imaging techniques as well as tissue sampling, others would advocate endoscopic examinations. Transvaginal ultrasound is a non-invasive method of imaging the endometrium but it may not be sensitive enough to detect endometrial polyps (Achiron *et al.,* 1995; Gupta *et al.,* 1996; Marconi *et al.,* 1997). Hysteroscopy may be valuable for detecting focal lesions but less so in general evaluation of endometrial hyperplasia. These issues require further research.

1.9 Literature review

Unlike systematic reviews, healthcare interventions of diagnostic tests are evolving. The various diagnostic tests to exclude endometrial cancer and hyperplasia evaluated in this thesis are: hysteroscopy, ultrasonography and endometrial sampling, as outlined above. The literature was searched using a combination of text words and MeSH representing these tests in MEDLINE database (1966-2001), EMBASE (1980-2001). The search was limited to human studies. It revealed a dearth of existing reviews. Only two meta-analyses were identified: one on the accuracy of endovaginal ultrasound by (Smith-Bindman *et al.,* 1998) and another one on the accuracy of Pipelle endometrial sampling by (Dijkhuizen *et al.,* 2000). Four more systematic reviews are now published but were initially obtained by direct personal contact with the authors as these reviews were still in press at the time of writing this thesis (Clark 2002a; Clark 2002b; Clark *et al.,* 2001; Gupta 2002). Table 1.1 summarises the conclusions of these six meta-analyses/systematic reviews.

Table 1.1: Conclusions of reviews on accuracy of tests for menstrual disorders

Authors	Meta-analyses/ systematic reviews	Summary quantitative results	Conclusions
(Smith-Bindman *et al.,* 1998)	Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. A meta-Analysis	Using a 5-mm threshold to define abnormal ET, 96% (95% CI 94-98%) of women with cancer had an abnormal scan, whereas 92% (95% CI, 90-93%) of women with endometrial disease (cancer, polyp, or atypical hyperplasia) had an abnormal result. This did not vary by HRT use. A woman with PMB and a 10% pre-test probability of endometrial cancer, has a 1% post-test probability following a normal scan	Endovaginal ultrasound has a high sensitivity for detecting endometrial cancer and other endometrial disease and can reliably identify postmenopausal women with vaginal bleeding who are highly unlikely to have significant endometrial disease so that endometrial sampling may be unnecessary
(Dijkhuizen *et al.,* 2000)	The accuracy of endometrial sampling in the diagnosis of patients with endometrial carcinoma and hyperplasia: a meta-analysis	In both postmenopausal and premenopausal women, the Pipelle was the best device, with detection rates of 99. 6% and 91%, respectively. For the detection of atypical hyperplasia, again, the Pipelle was the most sensitive technique with a sensitivity of 81%.	Endometrial biopsy with the Pipelle is superior to other endometrial techniques in the detection of endometrial carcinoma and atypical hyperplasia. The accuracy of the Pipelle is higher in postmenopausal women compared with premenopausal women.
(Clark *et al.,* 2001)	Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial hyperplasia. A systematic quantitative review	A positive test on outpatient biopsy diagnosed endometrial hyperplasia with a pooled LR of 12.0 (95% CI 7.8-18.6) while a negative test had a pooled LR of 0.2 (95% CI 0.1-0.3). With a positive test result, the post-test probability of endometrial hyperplasia was 57.7% (95% CI 41.1%-72.7%) while it was 2.2% (95% CI 0.9%-4.1%) with a negative test	Outpatient endometrial biopsy has modest accuracy in diagnosing endometrial hyperplasia.
(Clark 2002a)	Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: a systematic quantitative review	A positive test result on outpatient biopsy diagnosed endometrial cancer with a pooled LR of 66.48 (95% CI 30.04-147.13) while a negative test result had a pooled LR of 0.14 (95% CI 0.08-0.27). The post-test probability of endometrial cancer was 81.7% (95% CI 59.7%-92.9%) for a positive test and 0.9% (95% CI 0.4%-2.4%) for a negative test.	Outpatient endometrial biopsy has a high overall accuracy in diagnosing endometrial cancer when an adequate specimen is obtained. A positive test result is more accurate for ruling in disease than a negative test result is for ruling it out.
(Gupta 2002)	Ultrasonographic endometrial thickness for diagnosing endometrial pathology in women with postmenopausal bleeding. A meta-analysis	Using the pooled estimates, a positive test result raised the probability of carcinoma from 14.0% (95% CI 13.3-14.7) to 31.3% (95% CI 26.1-36.3), while a negative test reduced it to 2.5% (95% CI 0.9-6.4)	Ultrasound measurement of endometrial thickness alone, using the evidence from the best quality studies cannot be used to rule *in* presence of endometrial pathology accurately in women with postmenopausal bleeding. A negative result at ≤ 5 mm cut-off level measuring both endometrial layers rules *out* endometrial pathology with good certainty
(Clark 2002b)	Accuracy of hysteroscopy in the diagnosis of endometrial cancer and hyperplasia: a systematic quantitative review	The pre-test probability of endometrial cancer was 3.9% (95% CI 3.7-4.2%). A positive hysteroscopy(pooled LR 62.8, 95% CI 52.8-74.6) raised the probability to 71.8% (95% CI 67.0-76.6%), while a negative hysteroscopy (pooled LR 0.15, 95% CI 0.13-0.18) reduced the probability to 0.6 % (95% CI 0.5-0.8%). The 10.6% (95% CI 10.2-11.0%) pre-test probability of endometrial disease was raised to 55.1% (95% CI 52.6%-57.8%) with a positive hysteroscopy (pooled LR 10.4, 95% CI 9.7-11.1) while it was reduced to 2.8% (95% CI 2.5-3.0%) with a negative hysteroscopy (pooled LR 0.24, 95% CI 0.22-0.25	The diagnostic accuracy of hysteroscopy is high for endometrial cancer but only moderate for endometrial disease (cancer and/or hyperplasia).

A systematic review is one in which bias has been reduced by the systematic identification, appraisal, synthesis, and, if relevant, statistical aggregation of all relevant studies on a specific topic according to a predetermined and explicit method. The Quality of Reporting of Meta-analyses (QUOROM) conference (Moher ***et al.,*** 1999) addressed standards for improving the quality of reporting of meta-analyses of clinical randomised controlled trials and set items that they thought should be included in a checklist. Their checklist suggests that investigators explicitly describe all search strategies used to locate and appraise articles for inclusion in a meta-analysis. This includes description of the characteristics of primary studies; quantitative data synthesis; reliability and issues related to internal validity (Moher ***et al.,*** 1999); and clinical implications related to external validity (or generalisability). Table 1.2 summarises the assessment of the six identified meta-analyses/systematic reviews according to the QUOROM checklist.

Table 1.2: Quality assessment of reviews on accuracy of tests for menstrual disorders adopted from the QUOROM checklist

Study Feature	Quality assessment	(Smith-Bindman wt al., 1998)	(Dijkhuizen *et al.*, 2000)	(Clark *et al.*, 2001)	(Clark 2002a)	(Gupta 2002)	(Clark 2002b)
Title	**Identify the report as a meta-analysis/ Systematic review**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Abstract	Use a structured format	Yes	Yes	Yes	Yes	No	Yes
Objective	The question explicitly		Yes	Yes	Yes	Yes	Yes
Data sources	The databases & other sources	Yes	Yes	Yes	Yes	Yes	Yes
Review methods	The selection criteria (population, intervention, outcome/ study design)	Yes	Yes	Yes	Yes		Yes
	Methods for validity assessment, data abstraction,	Yes	Yes	Yes	Yes	Yes	Yes
	and quantitative data synthesis in sufficient detail to permit replication	Yes	Yes	Yes	Yes	Yes	Yes
Results and conclusion	Characteristics of studies included & excluded; Patient selection	Yes	Yes	Yes	Yes	Yes	Yes
	Qualitative and quantitative findings and subgroup analyses		Yes	Yes		Yes	Yes
Introduction	Describe						
	The explicit problem, the rationale for the intervention, and rationale for review	Yes	Yes	Yes	Yes	Yes	Yes
Methods							
Searching	The information sources in details	Yes	Yes	Yes	Yes	Yes	Yes
Selection	The inclusion and exclusion criteria	Yes	No	Yes	Yes	Yes	Yes
Validity assessment	The criteria & process used	Yes	Yes	Yes	Yes	Yes	Yes
Data abstraction	The process used	Yes	Yes	Yes	Yes	Yes	Yes
Study characteristics	The type of study design	Yes	Yes	Yes	Yes	Yes	Yes
Quantitative data Synthesis	The principal measures of effect (RR, CI) handling the missing data, heterogeneity assessment, subgroup analysis and publication bias	_	_	Yes	Yes	Yes	Yes
Results							
Trial flow	Meta-analysis profile	No	No	Yes	Yes	Yes	Yes
Study characteristics	Present descriptive data for each trial	Yes	Yes	Yes	Yes	Yes	Yes
Quantitative data Synthesis	Validity assessment present simple summary results	Yes	Yes	Yes	Yes	Yes	Yes
Discussion	Summarise key findings, describe potential biases, suggest a future research agenda	_	_	Yes	Yes	Yes	Yes

The existing reviews and meta-analyses only analysed data at study level. More intricate analysis requires the use of raw data and this is only possible with primary studies or with meta-analysis of individual patient data. This option needs further exploration.

There remain many issues not dealt with in these reviews. For instance the prevalence of malignancy among endometrial polyps, and the significance of insufficient sample on endometrial biopsy in relation to endometrial atrophy. These issues have been raised and addressed through primary studies in this thesis.

1.10 Endometrial polyps

The aetiology and pathogenesis of endometrial polyps are not fully understood but polyps are believed to be a risk factor for endometrial cancer (Maja Jr 1996). Only a few studies have looked at the pathological significance of endometrial polyps (Maja 1996; Maja Jr 1996). However, the issue of the pathological significance of endometrial polyps has seldom been looked at in controlled studies.

1.11 Insufficient sample

In the ambulatory investigation of abnormal uterine bleeding, an insufficient sample is obtained from outpatient endometrial biopsy in 25-33% cases (Antoni *et al.,* 1997; Gordon, Westgate *et al.,* 1999). Insufficient tissue has been reported in 12.8% of patients having Pipelle biopsy (Stovall *et al.,* 1991).

Generally, an insufficient sample is not believed to be a cause for concern because failure to obtain an endometrial specimen from a devise correctly positioned within the uterine cavity is considered to be an assurance that no significant intrauterine pathology is present (Ben Baruch ***et al.,*** 1994). However, this issue has not so far been exposed in a multivariable analysis.

1.12 Questions addressed in this thesis

With the background highlighted above this thesis raises and addresses the following research questions:

This thesis addresses the following research questions:

1. What is the accuracy of outpatient miniature hysteroscopy in the identification of premalignant and malignant endometrial lesions?
2. What is the accuracy of ultrasound scan in the identification of premalignant and malignant endometrial lesions?
3. What is the relative significance of hysteroscopic and ultrasonographic evidence of endometrial atrophy in relation to insufficient sample on outpatient endometrial biopsy?
4. What is the risk of premalignant and malignant pathology among endometrial polyps?
5. What is the significance of various risk factors associated with endometrial polyps?
6. What is the feasibility of multivariable analysis to evaluate combinations of diagnostic tests in the diagnosis of endometrial disease?

CHAPTER 2

THE DIAGNOSTIC ACCURACY OF OUTPATIENT MINATURE HYSTEROSCOPY IN PREDICTING PREMALIGNANT AND MALIGNANT ENDOMETRIAL LESIONS

2.1 Introduction

Menorrhagia, unscheduled bleeding on hormone replacement therapy and postmenopausal bleeding are common gynecological problems (Coulter 1995; Spencer 1999). As highlighted in chapter 1, the main aim of investigations for abnormal uterine bleeding is to exclude endometrial carcinoma and its precursor, endometrial hyperplasia either complex or atypical. This is because endometrial cancer is associated with abnormal uterine bleeding in over 90% of cases (Mencaglia 1995). Traditionally, abnormal uterine bleeding has been investigated with dilatation and curettage but now there is a trend towards minimally invasive investigations utilizing outpatient hysteroscopy, ultrasound scan and endometrial biopsy. However, the research –base underpinning these technologies is not as sound as one might think. The accuracy of various diagnostic tests in predicting premalignant and malignant endometrial lesions was not yet fully established by methodologically sound research at the time of conducting this original study (Al-Azzawi 1996; Bender ***et al.,*** 1997; Cicinelli 1993; Haller 1996; Mencaglia 1995). A recent review on this topic provided in chapter 1 (Clark 2002b) was not yet published at the time of writing. Despite the evidence from this review, at the present time, the evaluation of endometrial abnormalities remains a subject of a continuing debate. Although reviews of literature provide summary estimates of accuracy for various tests from analysis of data at study level, a detailed exploration using primary data is essential to inform the debate about the value of tests used in abnormal uterine bleeding.

A combination of ultrasound imaging, endoscopy and tissue sampling methods remain in common use to help minimise diagnostic error. Hysteroscopy has been used for direct endometrial visualisation to establish diagnosis of intrauterine pathology. Improvements in instrumentation have led to the development of small diameter endoscopes with improved quality and this has meant that hysteroscopy can be performed on an outpatient basis (de Jong ***et al.,*** 1990; Dounes 1993; Finikiotis 1994; Gimpleson 1984; Siegler 1995). However, endoscopic interpretation has not always been validated against histological findings. Studies comparing hysteroscopy with traditional blind dilatation and curettage have suggested the superiority of the former (Gimpleson 1998; Loffer 1989; Saidi 1997; Saidi ***et al.,*** 1997; Towbin 1996; Uno ***et al.,*** 1995) but these data generally relate to benign disease. Therefore, the results of these studies cannot be extrapolated with confidence to the hysteroscopic diagnosis of premalignant and malignant endometrial lesions.

There is a concern that with improvements in the quality of hysteroscopic technology a high degree of diagnostic accuracy is being assumed rather than established. The literature summarised in the review in chapter 1 revealed several relevant studies (Al-Azzawi 1996; Ben Yehuda ***et al.,*** 1998; Gimpleson 1998; Loffer 1989; Saidi 1997; Towbin 1996; Uno LH ***et al.,*** 1995) which evaluated the test performance characteristics of hysteroscopy, but they were generally of poor quality (Ben Yehuda ***et al.,*** 1998; Gimpleson 1998; Loffer 1989; Saidi 1997; Towbin 1996; Uno ***et al.,*** 1995). This background prompted this study and it addresses the first research question of this thesis: What is the accuracy of outpatient miniature hysteroscopy in the identification of premalignant and malignant endometrial lesions in women with abnormal uterine bleeding using histological assessment as a gold standard for the diagnosis of premalignant hyperplasia and carcinoma?

2.2 Ethical approval

The studies reported in this thesis were considered by the clinical governance department and the Professor in charge at the time (Professor John Newton). It was felt that it was not essential to obtain approval from the ethics committee. Data obtained from or about the subjects was treated as confidential and analysed anonymously. This was in compliance with the Declaration of Helsinki about Ethical Principles for Medical Research involving human subjects (Press Centre of the World Medical Association 1996). However the ethical principles were expanded to include research involving human data in the year 2000 (52nd WMA General Assembly 2000). If the research was conducted today, it would require formal approval from Research Ethical Committee. However, in this thesis, the data collection and analysis were performed before the new principles in Research Ethics were established in the year 2000. Therefore, the ethical basis and the procedures appropriate at the time of conducting this work had been followed and no ethical principles were violated. Letters from the Chief of Research and Development Committee and from the Clinical Governance Manager, confirming this is attached in Appendix.

2.3 Methods

To address the above question a prospective cohort study was conducted in a rapid access 'One-stop' outpatient service. Patients were referred to the service from gynaecology outpatient clinics or directly from general practice for investigation of abnormal uterine bleeding (menorrhagia, irregular menstrual loss, postmenopausal bleeding and unscheduled bleeding on tamoxifen or hormone replacement therapy). After consultation and counselling, a physical examination was performed, followed by hysteroscopy and outpatient endometrial biopsy. If the hysteroscopic view was poor or if the biopsy was not obtained, inpatient hysteroscopy and curettage under general anaesthesia were carried out. The outpatient hysteroscopy was performed after placing Instillagel (Farco-Pharma, Cologne, Germany) into the cervical canal via a long vaginal quill to achieve analgesia. For uterine

distension, normal saline solution was used with a syringe driven manually. The endoscopic equipment included 1.2 mm hysteroscope with 2.5 mm rigid sheath (K. Storz, Tuttlingen, Germany), a 175-watt xenon lamp light source and light cable, a video camera system and monitor. Hysteroscopy was performed by a consultant or by a senior minimal access surgery trainee. The depth of mucosa was evaluated by applying light pressure of the top of the endoscope against the endometrial surface. This evaluation was rather a subjective one based on the hysteroscopist's experience. Smooth thin endometrium without vascularisation was considered normal/benign (Ben Yehuda ***et al.,*** 1998; Cicinelli 1993). Features of increased endometrial thickness, abnormal vascularisation, polypoid formations, mamillations and cerebroid irregularities were considered suspicious lesions (Ben Yehuda ***et al.,*** 1998; Cicinelli 1993; Stovall 1991). Suspicious features associated with irregular polylobular, friable excrescences with necrosis or bleeding were considered diagnostic of premalignant/malignant endometrial lesions (Cicinelli 1993). In this manner the hysteroscopic features were classified as *negative* (normal/benign), *intermediate* (suspicious lesions) or *positive* (diagnostic of premalignant/malignant lesions).

The endometrial biopsies were performed using the Pipelle sampling device (Jaeschke 1994) (Laboratoire CCD, Paris, France). The Pipelle was blindly introduced in the uterine cavity to the fundus and the piston was withdrawn to the full length of the instrument, thereby creating negative pressure. As the instrument was withdrawn it was rotated drawing the tissue sample into the lumen. Histologic diagnoses were classified as *benign* (secretory/proliferative endometrium, benign endometrial polyp or simple hyperplasia), *premalignant* (complex or atypical hyperplasia) or *malignant* (carcinoma) based on the pathology report.

To estimate diagnostic accuracy, independent blind comparison of hysteroscopic findings (negative, intermediate or positive) was carried out with the histologic diagnoses (benign, premalignant or

malignant). Initially, the analysis was performed by collapsing the figures into binary data to generate a 2x2 table with hysteroscopic features (negative or intermediate/positive) in rows and histologic diagnosis (benign or premalignant/malignant) in columns. These data were used to calculate sensitivity, specificity and predictive values along with their 95% confidence intervals (CI).

Subsequent analysis was performed to delineate the significance of the hysteroscopic findings using multilevel tables. The accuracy of the various hysteroscopic features (negative, intermediate and positive) was statistically computed using likelihood ratio (LR) as the measure of predicting histologic diagnosis (Deeks 1996; Owen 1998). The imprecision around the LR values were estimated using 95% CI. The LR represented the ratio of the probability of a positive (or negative) test result in women with premalignant/malignant lesion to the probability of the same test result in women with benign endometrial histology. The LR indicated by how much a given hysteroscopic finding will raise or lower the probability of having premalignant/malignant lesion. With a positive hysteroscopy, an LR >1 increased the probability that premalignant/malignant lesion was present. The greater the LR, the larger the increase in the probability of premalignant/malignant lesion and the more clinically useful the test result. With a negative hysteroscopy, an LR <1 decreased the probability that premalignant/malignant lesion was present. The smaller the LR, the larger the decrease and the more clinically useful the test result.

Based on the previously validated analytic methodology (Chien 2001; Department of clinical epidemiology and biostatistics 1981; Khan 2001), the implications of the LRs associated with hysteroscopic findings was examined for the different pre-test probabilities in premenopausal and postmenopausal women by using Bayes' theorem to generate post-test probabilities as shown below:

Pre-test probability = prevalence of premalignant/malignant lesion

Pre-test odds = pre-test probability/ (1 – pre-test probability)

Post-test odds = likelihood ratio x pre-test odds

Post-test probability = post-test odds/ (1 + post-test odds)

An estimate of pre-test probability was obtained by calculating the prevalence of premalignant/malignant lesion in the different population sub-groups.

2.4 Results

During the period from November 1996 to December 1997, there were 248 patients who all had mini-hysteroscopy as well as histologic assessment of the endometrium. In 51/248 patients, the histology was obtained on inpatient assessment while in the rest the histology was obtained on outpatient assessment. The mean age of the patients was 50.6 years (range: 23-94) and the mean parity was 2.4 (range: 0-9) and there was 17 nulliparous subjects in this cohort of 248 patients. The reason for referral was postmenopausal bleeding in 112 (45%) patients while the rest 136 (55%) had premenopausal menstrual problems. Hormone replacement therapy was used by 73 (29.4%) women.

Outpatient hysteroscopy showed negative features in 228 patients, intermediate features in 15 and positive features in five patients. However, on histology, premalignant/malignant lesion of the endometrium was found in 18 cases. There were eight cases of endometrial cancer confirmed on histology, one in the premenopausal group. In addition, there was one case of cancer on histology that was considered as negative on hysteroscopy as it was an extended primary cervical cancer. The reason for its exclusion was that the objective of the study was to establish the accuracy of hysteroscopy in predicting endometrial cancer/hyperplasia. Therefore this case was not included in

the analysis. Of the 10 cases of complex or atypical endometrial hyperplasia, three were in premenopausal and seven in postmenopausal women.

The discordance between visual inspection of endometrial cavity by hysteroscopy and histologic diagnosis is shown in Table 2.1. Of the 228 patients classified as having normal/benign endometrium hysteroscopically, 12 were found to have premalignant or malignant lesion histologically. However, when the hysteroscopy was intermediate or positive, 14/20 turned out to have benign endometrial histology. Thus, the sensitivity of hysteroscopy was 33.3% (95% CI 14.4-58.8), specificity was 93.9% (95%CI 89.8-96.5), positive predictive value was 30.3% (95%CI 12.8-54.3), and negative predictive value was 94.7% (95%CI 90.8-97.1).

Multilevel analysis using likelihood ratio revealed that for negative hysteroscopy the LR was 0.7 (95% CI 0.5-0.9). For intermediate features at hysteroscopy the LR was 1.9 (95% CI 0.5-6.6) and for positive hysteroscopy it was 51.1 (95% CI 7.9-326.9) (Table 2.1).

Table 2.1: Diagnostic value of hysteroscopy in prediction of endometrial premalignant/malignant lesion.

	Histologic Diagnosis+			
	Benign	**Premalignant**	**Malignant**	**Likelihood Ratio$**
Hysteroscopic Features*				
Negative	216	8	4	**0.7 (0.5-0.9)**
Intermediate	13	1	1	**1.9 (0.5-6.6)**
Positive	1	1	3	**51.1 (7.9-326.9)**

* Hysteroscopic features are classified as *negative* (normal/benign), *intermediate* (suspicious lesions) or *positive* (diagnostic of premalignant/malignant lesions).
+ Histologic diagnoses are classified as *benign* (secretory/proliferative endometrium, benign polyp or simple hyperplasia), *premalignant* (complex or atypical hyperplasia) or *malignant* (carcinoma) based on the pathology report.
$ The LR represent the ratio of the probability of a positive (intermediate or negative) test result in women with premalignant/malignant lesion to the probability of the same test result in women with benign endometrial histology. With a negative hysteroscopy, an LR <1 decreased the probability that premalignant/malignant lesion was present, whereas with a positive hysteroscopy, an LR >1 increased the probability that such a lesion was present. See the methods section for a full description.

To determine if hysteroscopy was more or less useful in particular risk groups, the implication of these LRs on changes in probability of premalignant/malignant lesion in various risk groups was shown in Table 2.2.

As shown in Table 2.2, in postmenopausal patients, the pre-test probability of premalignant/malignant lesion was 12.5%. The post-test probability of having these lesions fell to 9.1% if the hysteroscopic findings were negative. However, with intermediate features the post-test probability increased to 21.3% and with positive features the probability increased substantially to 87.9 %. In premenopausal patients, the pre-test probability of premalignant/malignant lesion was 3.0%. The post-test probability of having these lesions reduced to 2.1% if the hysteroscopic findings were negative. However, with intermediate features the post-test probability increased to 5.5%. With positive features of abnormality the probability increased to 61.2%. This meant that miniature hysteroscopy had limited value in low risk patients but it was likely to be of some use in a high-risk group of patients.

Table 2.2: Diagnostic value of hysteroscopy in prediction of endometrial premalignant/malignant lesion in postmenopausal and premenopausal women.

Pre-test probability of premalignant/malignant lesion*	Hysteroscopic Features+	Post-test probability of premalignant/malignant lesion
Postmenopausal women		
12.5% (14/112)	Negative	9.1%
	Intermediate	21.3%
	Positive	87.9%
Premenopausal women		
3.0% (4/136)	Negative	2.1%
	Intermediate	5.5%
	Positive	61.2%

* An estimate of pre-test probability is obtained by calculating the prevalence of premalignant/malignant lesion in the different population sub-groups.
+ Hysteroscopic features are classified as *negative* (normal/benign), *intermediate* (suspicious lesions) or *positive* (diagnostic of premalignant/malignant lesions). See methods section for full description.

2.5 Discussion

The proposed structure by Docherty and Smith (Docherty 1999) for the discussion of scientific papers is used in this chapter and all other chapters for discussion in the following order:

- Statement of principal findings
- Appraisal of quality of the study: Strengths and weaknesses of the study
- Relation to other studies considering differences in quality and results
- Meaning of the study: possible mechanisms and implications for clinicians or policymakers
- Unanswered questions and future research

2.5.1 Principal findings of the study

This study showed that negative or intermediate findings at hysteroscopy are not conclusive in the prediction of endometrial premalignant/malignant lesion. This meant that the probability of premalignant/malignant lesion was not particularly altered by the knowledge of normal hysteroscopic appearances. However, if the findings were definitely positive of abnormality, the LR was large and it substantially increased the post-test probability of confirming premalignant/malignant lesion. Furthermore this study showed that miniature hysteroscopy had limited value in low risk patients but it was likely to be of some use in a high-risk group of patients. As normal findings do not eliminate the need for endometrial sampling, miniature hysteroscopy is not a substitute for histopathological assessment in an outpatient diagnostic service.

2.5.2 Appraisal of quality of the study

The strength of this study and the credibility of its findings depend on the strength of the methodology. This is made clear by adhering to the checklist proposed by the STARD group (Standards for Reporting of Diagnostic Accuracy), the objective of which is to improve the

quality of reporting of studies on diagnostic accuracy in order to detect the potential for bias in the study (Table 2.3).

Table 2.3: Application of STARD checklist to the study reported in Chapter 2

Section and topic	Description of the criteria	Compliance of thesis
METHODS		
Participants:	Based on presenting symptoms, the participants had received the index test(s) and the reference standard?	Yes
Study population		
Setting/ Location		
Recruitment:		
Participant sampling:	Consecutive series of patients defined by selection criteria	Yes
Data collection:	Participants identified and data collected before the index test(s) and reference standards were performed (prospective study)	Yes
Reference standard	and its rationale	Yes
Test methods	Technical specification of given material and methods	Yes
	Definition and rationale for the units, cut-offs and/or categories of the results of the index test(s) and the reference standard	Yes
	The number, training and expertise of the persons (a) executing and (b) reading the index test(s) and the reference standard	Yes
	Methods for calculating measures of diagnostic accuracy or making comparisons, the statistical methods used to quantify uncertainty (e.g. 95% CI) and methods for calculating test reproducibility	Yes
RESULTS		
Participants	When study was done, including beginning and ending dates of recruitment	Yes
	Clinical and demographic characteristics (e.g. age, sex, spectrum of presenting symptom(s), co morbidity, current treatment(s), recruitment centre)	Yes
Test results	A cross tabulation of the results of the index test(s) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard	Yes
	Indeterminate results, missing responses and outliers of index test(s) stratified by reference standard result and how they were handled	Yes
Estimation	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% CI)	Yes
	Estimates of variability of diagnostic accuracy between subgroups of participants (Post vs premenopausal)	Yes

The criteria that distinguish a definitely useful clinical test from one that is of doubtful value are generally well established and have been embodied in the rules of critical appraisal of the medical literature (Chien 1997; Deeks 1996; Department of clinical epidemiology and biostatistics 1981; Greenhalgh 1997; Jaeschke ***et al.,*** 1994; Owen 1998). According to these rules, the new and more powerful concepts around LRs are preferable to the old-fashioned ideas of sensitivity and specificity in the assessment of the clinical value of a test. By applying the criteria for usefulness of a test based on LRs, a diagnostic test study can more meaningfully summarise the findings of the research, minimising the risk of erroneous inferences. For a negative test result LR <0.1 is regarded as definitely useful, 0.1-0.2 is regarded as moderately useful, 0.5-0.2 is regarded as slightly useful, and 0.5-1 is regarded as not at all useful. For a positive test result, LR >10 is regarded as definitely useful, LR 5-10 is regarded as moderately useful, LR 2-5 is regarded as slightly useful, and LR 1-2 is regarded as not at all useful. When applied to the study's results, these criteria indicate that a negative or intermediate hysteroscopy was not at all useful but a positive hysteroscopy was definitely useful.

2.5.3 Relation to other studies considering differences in quality and results

Is it surprising that an examination involving direct visualisation of the entire uterine cavity with excellent image quality and magnification has limited diagnostic value as shown in this study? To answer this question, the results based on more traditional concepts evaluating diagnostic accuracy of hysteroscopy were compared with this study. This study's analysis showed that hysteroscopy had a sensitivity of 33.3% (95% CI 14.4-58.8) and specificity of 93.9% (95% CI 89.8-96.5) in the detection of premalignant/malignant lesions. Similarly, Haller *et al* quoted sensitivity of 50% and specificity of 100% for endometrial carcinoma and Uno *et* al (Uno LH ***et al.,*** 1995) reported sensitivity of 15.8% and specificity of 97.3% for endometrial hyperplasia. As a sensitive test when negative rules out disease, it is clear that the appropriate inference should be that a normal

hysteroscopy is poor in ruling out cancer or hyperplasia. Ben Yehuda *et al* (1998) also reached a conclusion of poor diagnostic performance for hysteroscopy in the prediction of hyperplasia and cancer, which concurs with my inferences.

With the lack of rigorous evidence in this field, this primary study was the original work that prompted the systematic review within our group (Clark 2002b). The results of this systematic review confirmed the findings of this primary study. It revealed that for endometrial cancer a positive hysteroscopy had a pooled LR of 62.8 (95% CI 52.8-74.6), while a negative hysteroscopy had a pooled LR of 0.15 (95% CI 0.13-0.18). The review concluded that the diagnostic accuracy of hysteroscopy is high for endometrial cancer but only moderate for endometrial disease defined as cancer and/or hyperplasia, an approach that has been used in an earlier review of endometrial ultrasound (Smith-Bindman *et al.,* 1998).

2.5.4 Practical implications for clinicians

Hysteroscopy is highly accurate and thereby clinically useful in diagnosing endometrial cancer in women with abnormal uterine bleeding. Endometrial sampling and/or ultrasonography should be considered with negative hysteroscopy, as negative hysteroscopy is not conclusive for excluding malignancy.

2.5.5 Implications on research

- Guideline or national consensus on how to investigate menstrual disorders/abnormal uterine bleeding is needed. Such a guideline was recently published in Scotland (the Scottish Intercollegiate Guideline Network 2002).
- Systematic reviews using Individual Patients Data or large primary studies are needed to evaluate accuracy of combinations of tests used in ambulatory assessment of menstrual

disorders.

- An estimate of the effectiveness and cost effectiveness estimation of the available diagnostic tools (hysteroscopy /ultrasound /biopsy) are required.

CHAPTER 3

THE DIAGNOSTIC ACCURACY OF ULTRASOUND SCAN IN PREDICTING ENDOMETRIAL HYPERPLASIA AND CANCER IN POSTMENOPAUSAL BLEEDING

3.1.1 Introduction

The rate of endometrial cancer and hyperplasia in cases of postmenopausal bleeding is over 10-15% and the investigations for this condition are aimed at identifying these pathological conditions (Gupta ***et al.,*** 1996; Mencaglia 1995) . As highlighted in chapter 1, postmenopausal bleeding has ttraditionally been investigated with dilatation and curettage but now there is a trend towards minimally invasive investigations utilizing ultrasound scan, outpatient hysteroscopy, and endometrial biopsy. However, at the time of conducting this primary study, the accuracy of various diagnostic tests in predicting premalignant and malignant endometrial lesions was not yet fully established by methodologically sound research. The reviews in this field summarised in Chapter 1 (Smith-Bindman ***et al.,*** 1998; Gupta 2002) were not yet published. However, even given the evidence from these reviews, the evaluation of endometrial abnormalities remains a subject of continuing debate. Although reviews of literature provide summary estimates of accuracy for various tests from analysis of data at study level, a detailed exploration using primary data is essential to inform the debate about the value of tests used in postmenopausal bleeding.

A combination of ultrasound imaging, endoscopy and tissue sampling methods are employed to minimise diagnostic error in the assessment of postmenopausal bleeding. While some researchers consider hysteroscopy to be the standard of care (Dubimsky 1997), others have suggested that non-

invasive assessment using ultrasound may exclude endometrial pathology in postmenopausal women. Advances in technology have made possible the introduction of transvaginal ultrasound scanning using higher frequency ultrasound with greater proximity to the uterus. Hence, the endometrial-myometrial interface can be more clearly delineated than with the full bladder transabdominal technique.

The ultrasound endometrial echogenicity associated with the presence of glands and mucin is proportional to the growth and differentiation of the endometrium (Dorum 1993). Therefore, the endometrium of a normal atrophic uterus is thin, on average measuring 2.3mm by sonography (Fleisher 1996; Nasri 1989; Nasri 1991). Endometrial thickness (ET) test might be considered negative when the measurement is <5mm in women who are not receiving estrogen replacement therapy (Nasri 1989) and in these cases histological examination might not be needed (Granberg 1991). However, advanced carcinoma with minimal lesions is known to occur in cases without noticeable ultrasonic thickness of the endometrium (Dorum 1993; Gupta ***et al.,*** 1996). With this background, this study was carried out to address the following research question: What is the accuracy of ultrasound scan in the identification of premalignant and malignant endometrial lesions in women with postmenopausal bleeding using histological assessment as a gold standard for the diagnosis of endometrial cancer and hyperplasia?

3.1.2 The effects of hormone replacement therapy on the endometrium

There is evidence that unopposed oestrogen therapy may induce endometrial stimulation and increase the risk of endometrial hyperplasia, atypia, and the potential for endometrial carcinoma. Pickar *et al*'s review provides confirmation that the addition of progestogens administered either

cyclically or continuously, provides protection against endometrial hyperplasia (Pickar *et al*., 1998). The continuous presence of progestogens causes down-regulation of oestrogen and progestogens receptors and endometrial atrophy. The addition of progestogens may cause unwanted side effects such as unscheduled bleeding. However unscheduled bleeding is less likely under sequential than continuous therapy (Lethaby 2003).

One randomized, comparative study (Hanggi 1997) in postmenopausal women was designed to evaluate the endometrium under the effect of sequential HRT. It compared transvaginal ultrasonography immediately before Pipelle biopsy at baseline and subsequently after 12 and 24 months in those patients taking HRT and those who were not. The study found that after 24 months, endometrial thickness was increased in the HRT group comparing to the controls ($P < 0.001$). Moreover the study found that if endometrial thickness was <5 mm, the endometrial biopsy sample was either inactive/atrophic or insufficient for histopathological diagnosis. Hyperplastic or malignant changes were not reported.

The Nordic multicenter study (Granberg 1997) revealed that atrophy was found significantly more often in women without HRT. Hormonal effects on the endometrium were found significantly more often in women with HRT (sequential). Endometrial hyperplasia was found most commonly in women with oestrogen as compared to those with HRT (sequential) and without HRT. Endometrial cancer occurs most in women without HRT. Endometrial cancer did not occur in any woman (with oestrogen, sequential HRT, or without HRT) with an endometrial thickness of $\leq$4 mm.

It is in general true that the thicker the endometrium of a postmenopausal woman the higher the likelihood of endometrial cancer being present (Wolman 1998). The threshold value for deciding if there is a significant risk of cancer will depend, in part, on the “normal” endometrial thickness for

the relevant patient group. For example, women on sequential HRT tend to have thicker endometria, compared to those who have never used HRT, have not used any form of HRT for a year or more, or are using continuous combined HRT, and so a higher threshold needs to be used for identifying those with a significant probability of cancer (Scottish Intercollegiate Guidelines Network 2002).

3.2 Methods

A prospective cohort study was conducted in a rapid access 'One-stop' outpatient service in two large centres: Birmingham Minimal Access and Surgical Training centre (MAST), and Gynaecology Minimal Access Unit (GMAU), Ninewells Hospital, Dundee (see Appendix 2). The ethical considerations are covered in Section 2.2.

This study recruited 96 consecutive postmenopausal women, 44 referred to the MAST centre and 52 to the GMAU between November 1996 and December 1997. These patients were referred from gynaecology outpatient clinics or directly from general practice for the investigation of postmenopausal bleeding defined as vaginal bleeding at least six months after cessation of the menses. Contrary to chapter 2, this chapter's population included only women with postmenopausal bleeding because endometrial thickness on ultrasound scan is significant only in this group of women (postmenopausal), unlike hysteroscopy where it may be performed as an investigation for the entire spectrum of menstrual disorders in any age group.

After consultation and counselling, a physical examination was performed followed by transvaginal ultrasound scan of the pelvis and uterus using a portable scanner with an endovaginal 6.5 MHz transducer (Hitachi Sumi, Japan). Trained ultrasonographers carried out the scans. In order to reduce the inter-operator variability and be able to further generalise the results, the following

criteria were followed in both centres: the transducer was introduced into the posterior vaginal fornix, and the uterus was scanned longitudinally and transversely. The measurements of the endometrium were performed in the longitudinal planes, and scanning from cornua to cornua identified the thickest part. The measurement was performed between the two basal layers of the anterior and posterior uterine wall without including the measurements of any polyp and/or fluid that might be in the cavity (Dorum 1993; Gupta ***et al.,*** 1996; Karlsson 1995).

As a gold standard, the histopathologic diagnosis was used based on outpatient endometrial biopsies or inpatient curettage. Pipelle sampler was the device used to perform the endometrial biopsies (Laboratoire CCD, Paris, France) (Stovall 1991). The Pipelle was introduced in the uterine cavity to the fundus, and to create negative pressure the piston was withdrawn to the full length of the instrument. As the instrument was withdrawn it was rotated drawing the tissue sample into the lumen. Histologic diagnoses were classified as *benign* (secretory/proliferative endometrium, benign endometrial polyp), hyperplasia (simple, complex and atypical) equal to *premalignant* in Chapter 2, or *malignant* (carcinoma) based on the pathology report.

To estimate the diagnostic accuracy of endometrial thickness on ultrasound examination, the findings were blindly compared with the histologic diagnosis. Based on the work by Karlsson ***et al.,*** (1995), endometrium thickness $\geq$4.0mm was used as the cut-off level for abnormality. Initially, the analysis was performed by collapsing the figures into binary data to generate a 2x2 table with endometrial thicknesses ($\leq$4.0mm or >4.0mm) in columns and histologic diagnosis (benign or hyperplasia/cancer) in rows. These data were used to calculate sensitivity, specificity and predictive values along with their 95% CI. Subsequent analysis was performed to delineate the significance of the ultrasonographic endometrial thickness using multilevel tables. The accuracy of the various endometrial thicknesses ($\leq$4.0mm, 4.1-9.0mm, >9.0mm) was statistically computed using LR as the

measurement of predicting histologic abnormality (Deeks 1996; Jaeschke 1994). Setting these three cut-off values was arbitrary. The imprecision around the LR values was estimated using 95% CI. The descriptions of LR's are covered in Section 2.3.

Based on the previously validated analytical methodology (Khan 2001; Chien 2001; Owen 1998), the implications of the LRs associated with ultrasonographic endometrial thickness for the different pre-test probabilities in postmenopausal women were examined by using Bayes' theorem to generate post-test probabilities as shown in Section 2.3. An estimate of pre-test probability was obtained by calculating the prevalence of premalignant/malignant lesion in the study population.

3.3 Results

The mean age of the patients was 52.3 years (range: 42-94) and the mean parity was 2.4 (range: 0-9). Ultrasound scans showed endometrial thickness $\leq$4.0mm in 42 patients, endometrial thickness 4.1-9.0mm in 29 and endometrial thickness >9.0mm in 25 patients. However, on histology, endometrial hyperplasia/cancer was found in 14 cases. There were 11 cases of endometrial cancers, five in the MAST centre and six in the GMAU; and there were three cases of endometrial hyperplasia, one in the MAST centre and two in the GMAU. Not a single case of malignancy was found when the endometrial thickness was $\leq$4.0mm. The one case of hyperplasia that was found in this group was simple hyperplasia. The 11 cases of carcinoma had endometrial thickness >4.0mm.

Of the total of 96, 46 patients (47.9%) were on hormone replacement therapy (HRT) either continuous combined or sequential combined [32/44 (72.7%) in the MAST centre and 14/52 (26.9%) in the GMAU]. There was one case of cancer and two of hyperplasia among HRT users. In the MAST centre, one patient had hyperplasia and the endometrial thickness was 4mm in this case.

In the GMAU, one woman had cancer and the endometrial thickness was 7mm in this case, while another one had hyperplasia and the endometrial thickness in this case was 5mm.

The discordance between ultrasonographic endometrial thickness and histologic diagnosis is shown in Table 3.1. For detection of endometrial hyperplasia/carcinoma, endometrial thickness ≥4.0mm had a sensitivity of 92.9% (95%CI 64.2-99.6), specificity of 50% (95%CI 38.8-61.2), a positive predictive value of 24.1% (95%CI 13.9-37.9), and a negative predictive value of 97.6% (95%CI 85.9-99.9).

Multilevel analysis using likelihood ratio revealed that for endometrial thickness ≤4.0mm at ultrasound scan the LR was 0.14 (95%CI 0.02-0.64). For ET: 4.1-9.0mm at ultrasound scan the LR was 0.94 (95%CI 0.37-1.99) and for endometrial thickness >9.0mm it was 3.3 (95%CI 1.73-5.73) (Table 3.1).

The implication of these LRs on changes in probability of premalignant/malignant lesion is shown in Table 3.2. The table shows that in postmenopausal patients, the pre-test probability of endometrial hyperplasia/cancer was 15.0%. The post-test probability of having these lesions fell to 2.4% if endometrial thickness ≤4.0mm. However, with endometrial thickness 4.1-9.0mm the post-test probability changed to 14.2% and with endometrial thickness >9.0mm the probability increased to 36.8%.

Table 3.1: Diagnostic value of ultrasound scan in diagnosis of endometrial cancer and hyperplasia

	Histology	Benign	Hyperplasia	Carcinoma	Likelihood Ratio
Endometrial Thickness					
≤4.0mm		41	1	0	0.14 (0.02-0.64)
4.1-9.0mm		25	1	3	0.94 (0.37-1.99)
>9.0mm		16	1	8	3.3 (1.73-5.73)

+ Histologic diagnoses are classified as *benign* (secretory/proliferative endometrium, benign polyp), hyperplasia (simple, complex, or atypical) or *malignant* (carcinoma) based on the pathology report. $ The LR represents the ratio of the probability of a positive test result in women with endometrial cancer/hyperplasia to the probability of the same test result in women with benign endometrial histology. An LR <1 decreases the probability that endometrial cancer/hyperplasia is present, whereas an LR >1 increases the probability that such a lesion is present. See the methods section for a full description.

Table 3.2: Diagnostic value of ultrasound scan in diagnosis of endometrial cancer or hyperplasia

Pre-test probability of * endometrial hyperplasia/cancer	Endometrial thickness	Post-test probability of endometrial hyperplasia/cancer
15% (14/96)	≤4.0 mm	2.4 %
	4.1-9.0 mm	14.2 %
	>9.0 mm	36.8 %

* An estimate of pre-test probability is obtained by calculating the prevalence of endometrial hyperplasia/cancer in the study population.

3.4 Discussion

3.4.1 Principal findings of the study:

This study concluded that endometrial thickness at ultrasound scan is not conclusive in the prediction of endometrial hyperplasia/cancer but it is reliable in excluding such a lesion if endometrial thickness is ≤4.0mm although this finding does not totally eliminate the need for endometrial sampling as there was one case of simple hyperplasia found in this group. This meant that in women with postmenopausal bleeding, malignancy might be safely excluded if sonographic endometrial thickness was ≤4.0mm. However, the probability of endometrial hyperplasia/cancer is not particularly altered by the knowledge that endometrial thickness on scan was 4.1-9.0mm. For an endometrial thickness >9.0mm, the likelihood of endometrial hyperplasia/cancer is only slightly increased.

3.4.2 Appraisal of quality of the study

The strength of this study and the credibility of its findings depend on the strength of the methodology. This is made clear by adhering to the checklist proposed by the STARD group (Standards for Reporting of Diagnostic Accuracy), the objective of which is to improve the quality of reporting of studies on diagnostic accuracy allowing to detect the potential for bias in the study. (Table 3.3).

Table 3.3: Application of STARD checklist to the study reported in Chapter 3

Section and topic	Description of the criteria	Compliance of this study
METHODS		
Participants:	Based on presenting symptoms, the participants had received the index test(s) and the reference standard?	Yes
Study population		
Setting/ Location		
Recruitment:		
Participant sampling:	Consecutive series of patients defined by selection criteria	Yes
Data collection:	**P**articipants identified and data collected before the index test(s) and reference standards were performed (prospective study)	Yes
Reference standard	and its rationale	Yes
Test methods	Technical specification of Given material and methods	Yes
	Definition and rationale for the units, cut-offs and/or categories of the results of the index test(s) and the reference standard	Yes
	The number, training and expertise of the persons (a) executing and (b) reading the index test(s) and the reference standard	Yes
	Methods for calculating measures of diagnostic accuracy or making comparisons, the statistical methods used to quantify uncertainty (e.g. 95% CI) and methods for calculating test reproducibility	Yes
RESULTS		
Participants	When study was done, including beginning and ending dates of recruitment	Yes
	Clinical and demographic characteristics (e.g. age, sex, spectrum of presenting symptom(s), co morbidity, current treatment(s), recruitment centre)	Yes
Test results	A cross tabulation of the results of the index test(s) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard	Yes
	Indeterminate results, missing responses and outliers of index test(s) stratified by reference standard result and how they were handled	Yes
Estimation	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% CI)	Yes
	Estimates of variability of diagnostic accuracy between subgroups of participants (Post vs premenopausal)	Yes

In the rules of critical appraisal of the medical literature, the criteria that distinguish a definitely useful clinical test from one that is of doubtful value are discussed in Section 2.5.2.

This study may be criticised as being flawed as the analytic strategy did not perform separate statistics for the subgroup of women using HRT. The data on HRT use in postmenopausal bleeding was explored and revealed that despite half of the patients being on HRT (46/96), only one of the 11 patients of cancers occurred in HRT users. The endometrial thickness in this case was abnormal at 7mm. Keeping in mind that in this study, endometrial cancer did not occur in any women (with or without HRT) with an endometrial thickness of $\leq$4mm, it is safe to conclude that the findings about the prediction of malignant disease were not particularly influenced by HRT use. However, the results on premalignant disease were somewhat inconsistent with those of Granberg *et al.,* (1997) who reported that endometrial hyperplasia was found most commonly in women on oestriol in the endometrial thickness group 5-8mm (see 3.2.2). In two patients on HRT with endometrial hyperplasia in the thesis' study, the endometrial thickness was 4mm and 5mm.

3.4.3 Relation to other studies considering differences in quality and results

Is it surprising that an examination involving excellent image quality such as ultrasound scan has limited diagnostic value in our study? To answer this question, these results were compared with other studies that were generally based on more traditional concepts of evaluating diagnostic accuracy. Based on their analyses, Granberg ***et al.,*** (1991) and Nasri ***et al.,*** (1989) have suggested that an endometrial thickness of 5mm is an appropriate cut off level for conservative management of patients with postmenopausal bleeding while Giusa- Chiferi ***et al.,*** (1996) recommended no need for anatomopathologic tests when the thickness is <3mm. However, (Dorum ***et al.,*** 1993) quoting sensitivity of 80% and specificity of 60%, concluded that scan was not accurate enough to replace histologic examination of the endometrium for detecting endometrial malignancy. In their

interesting study Kufahl (1997) found that at a cut-off limit of 4mm, the sensitivity of the ultrasound was 90.3%, the specificity 24.8%, the positive predictive value 21.4% and the negative predictive value 91.9%. One endometrial cancer, one atypical and one complex hyperplasia were missed.

In their large multicentre study Karlsson *et al.,* (1995) concluded that the risk of finding pathologic endometrium at curettage is 5.5% when the endometrium is ≤4.0mm as measured by transvaginal ultrasonography. Thus in women with postmenopausal bleeding and an endometrium ≤4.0mm, it would seem justified to refrain from curettage. Dubimsky ***et al.,*** (1997) reported that transvaginal ultrasound appears to be more sensitive than endometrial biopsy for the detection of endometrial abnormalities. Saidi ***et al.,*** (1997) have quoted a sensitivity of 95% and a specificity of 65% for transvaginal ultrasonography in detecting abnormal findings.

Recently, there has been a metaanalysis carried out by Smith-Bindman *et al.,* (1998) to determine the accuracy of transvaginal scan in detecting endometrial disease in postmenopausal women with vaginal bleeding according to hormone replacement use. The overall summary mean weighted estimates of sensitivity and specificity were calculated for thresholds of endometrial thickness from 3 to 10 mm. Using a 5-mm threshold to define abnormal endometrial thickening, 96% (95% CI 94-98%) of women with cancer had an abnormal scan result, whereas 92% (95% CI 90-93%) of women with endometrial disease (cancer, polyp, or atypical hyperplasia) had an abnormal result. This did not vary by hormone replacement use. However, the number of women with normal histology who had an abnormal scan result did vary by hormone replacement use. In women who were not using hormone replacement therapy, 593 (8%) with normal histological findings had an abnormal ultrasound result (specificity, 92%; 95% CI 90-94%), whereas 1544 (23%) using hormone

replacement therapy had an abnormal scan result (specificity, 77%; 95% CI 75-79%). For a postmenopausal woman with vaginal bleeding with a 10% pre-test probability of endometrial cancer, her probability of cancer is 1% following a normal scan result. The conclusion from this meta analysis confirmed my findings that transvaginal ultrasound has a high sensitivity for ruling out endometrial cancer and other endometrial disease and can reliably identify postmenopausal women with vaginal bleeding who are highly unlikely to have significant endometrial disease. Therefore endometrial sampling may be unnecessary.

The primary study reported in this Chapter has been the original work that prompted the systematic review within our group jointly with the Ninewells centre. The results of this systematic review confirmed the findings of my primary study. Using the pooled estimates, a positive test result raised the probability of carcinoma to 31.3% (95% CI 26.1 - 36.3), while a negative test reduced it to 2.5% (95% CI 0.9 - 6.4). The review concluded that ultrasound measurement of endometrial thickness alone, using the evidence from the best quality studies couldn't be used to rule *in* presence of endometrial pathology accurately in women with postmenopausal bleeding. However, a negative result at ≤ 5mm cut-off level measuring both endometrial layers rules *out* endometrial pathology with good certainty (Gupta 2002).

Some reviewers prefer to give the rationale (by giving Receiver Operating Characteristic curve analysis) on how the various cut-off levels for normal measurement of endometrial thickness was arrived at for each of the above mentioned studies, however, most of these study are not powered enough to use the ROC. Moreover, in my opinion, the ROC lack the clinical usefulness as clinicians find it difficult to interpret.

3.4.4 Practical implications for clinicians

Endometrial sampling may safely be omitted in women with postmenopausal bleeding, if sonographic endometrial thickness is ≤4.0mm provided the whole cavity is adequately examined and it is regular.

Improvements in the technology might improve the accuracy of sonography in detecting endometrial hyperplasia/cancer in the future. It has been suggested that transvaginal colour and pulsed doppler sonography may help in improving differentiation of malignant from benign endometrial pathologic conditions in postmenopausal women (Aleem ***et al.,*** 1995).

3.4.5 Implications on research

- National guideline or national consensus on how to investigate menstrual disorders/abnormal uterine bleeding is needed. Such a guideline was although published in Scotland (the Scottish Intercollegiate Guideline Network 2002).
- Systematic reviews using Individual Patients Data of relevant or large primary studies are needed to evaluate the accuracy of combinations of tests used in ambulatory assessment of menstrual disorders (see Chapter 7).
- The effectiveness and cost effectiveness estimation of the available diagnostic tools (hysteroscopy /ultrasound /biopsy) are required.

CHAPTER 4

CONTROLLED ANALYSIS OF FACTORS ASSOCIATED WITH INSUFFICIENT SAMPLE ON OUTPATIENT ENDOMETRIAL BIOPSY

4.1 Introduction

In ambulatory investigation of abnormal uterine bleeding, an insufficient sample is obtained from outpatient endometrial biopsy in 25-33% cases (Antoni ***et al.,*** 1997; Gordon ***et al.,*** 1999). Generally, an insufficient sample is not believed to be a cause for concern as failure to obtain an endometrial specimen from a devise correctly positioned within the uterine cavity is considered to be an assurance that no significant intrauterine pathology is present (Ben Baruch ***et al.,*** 1994). Do we have the confidence that insufficient sample is benign in women with abnormal uterine bleeding? One study supporting this view (Choo 1985) showed that in 42% of cases with hysteroscopically atrophic endometrium, no tissue was obtained on traditional dilatation and curettage for histological assessment. Such data provide a plausible, but not scientific, evidence of association because of a lack of a comparison with a hysteroscopically normal control group. Moreover, such data also suffers from lack of an adjustment for confounding variables like age, menopausal status and ultrasonographic endometrial thickness. Quite apart from this methodological debate, information obtained from inpatient dilatation and curettage can hardly be considered generalisable to outpatient biopsy devices.

To have the confidence that insufficient sample on outpatient biopsy is benign in women with abnormal uterine bleeding; the following question should be answered:
What is the relative significance of hysteroscopic and ultrasonographic evidence of endometrial atrophy in relation to insufficient sample on outpatient endometrial biopsy?

4.2 Methods

To answer the above question, a controlled statistical analysis was conducted to delineate the relative significance of factors associated with insufficient sample on outpatient endometrial biopsy.

The database of 248 consecutive patients with abnormal uterine bleeding investigated in the One-stop clinic from November 1996 to December 1997 was used (Bakour 1999). Ethical considerations were discussed in Section 2.2.

In the clinic, a combination of pelvic ultrasound scan, outpatient hysteroscopy, and Pipelle endometrial sampling were performed. Pelvic transvaginal ultrasonography was performed using a portable scanner (Hitachi Sumi, Japan) with an endovaginal 6.5 MHz transducer. Endometrial thickness was measured in mm including both layers of the endometrium. Outpatient hysteroscopy was performed using 1.2mm micro-hysteroscope with 2.5mm outer rigid sheath (Karl Storz, Tuttlingen, Germany). Assessing the depth of mucosa by applying light pressure of the top of the endoscope, hysteroscopic appearances of thin smooth non-vascular endometrium were considered to be atrophic. Outpatient endometrial biopsies were performed using Pipelle sampling device (Laboratoire CCD, Paris, France). Of the 248 patients, 74 (29.8%) had insufficient endometrial sample on pathology report. All 74 women were followed using the hospital clinical computer system for any further admission episodes related to abnormal uterine bleeding over a period of two years. Inpatient assessment with sampling was performed on those with recurrent symptoms. None of these patients had endometrial cancer or hyperplasia.

To delineate the relative significance of various factors (independent variables) on sufficiency of outpatient endometrial biopsy (binary dependent variable), a multivariable logistic regression model (Khan 1999) was built. The primary null hypothesis we wanted to test was, "the value of the β-

coefficient for the various risk factors in the logistic model is zero or its exp (β) is 1.0. To obtain a valid estimate of the association between endometrial atrophy (on hysteroscopy and on ultrasonography) and insufficient endometrial specimen, the analysis was adjusted for the confounding effects of age, menopausal status, parity and use of hormone therapy by including them as independent variables in the multivariable logistic regression analysis. In the model building process, the significance of the independent variables was tested individually and in various combinations. Variables were tested for interaction and they were analysed using both nominal and continuous scales of measurement. The assumption of linearity of association was checked for the different independent variables. Cases with missing data were excluded from the multivariable analysis. The model that best predicted the odds of insufficient sample on endometrial biopsy with the lowest deviance and highest goodness of fit was selected. The Hosmer-Lemeshow test using the "deciles of risk" method was used to assess the goodness of fit.

4.3 Results

The characteristics of the women attending the One-stop clinic were shown in Table 4.1.

Table 4.1: Characteristics of women attending outpatient One-stop hysteroscopy clinic

Patients' Characteristics	Sufficient sample on endometrial biopsy N=174	Insufficient sample on endometrial biopsy N=74
Age - years		
Mean (95% CI)	50.7 (0.26-0.38)	50.6 (0.61-0.73)
Parity		
Mean (95% CI)	2.4 (0.26-0.38)	2.4 (0.61-0.73)
Post menopause		
n % (95% CI)	56 32.2% (25.3-39.7%)	56 75.6% (64.3-84.9%)
Hormone Replacement Therapy	44	29
n % (95% CI)	25.2% (19.0-32.4%)	39.2% (28.0-51.2%)
Cancer		
n % (95% CI)	8 4.6% (2.0-8.9%)	0 0% (0-4.8%)

Univariable regression (Table 4.2) revealed that atrophic endometrium on hysteroscopy increased the odds of having insufficient sample on endometrial biopsy (OR 5.53, p=0.001) while endometrial thickness above 5mm on ultrasonography decreased the odds (OR 0.15, p=0.001). Increasing patient's age (OR 1.04, p=0.001) and postmenopausal status (OR 4.49, p=0.001) were also associated with insufficiency of outpatient sampling in unadjusted analysis. When these variables were combined in a multivariable regression analysis to control for confounding, only endometrial atrophy on hysteroscopy (OR 4.79, p=0.04) and endometrial thickness above 5mm on ultrasonography (OR 0.19, p=0.001) remained significant in predicting an insufficient sample on outpatient endometrial biopsy. In a similar but separate multivariable model using endometrial thickness as a continuous variable we found that the odds of insufficient sample decreased with each 1mm increase in endometrial thickness (OR 0.88, 95% CI 0.78-0.99, p=0.03).

Table 4.2: Logistic regression analysis to determine the relationship between various factors associated with insufficient specimen on outpatient endometrial biopsy

Risk Factor	Univariable Regression		Multivariable Regression*	
	Odds Ratio (95% Confidence Intervals)	**P value**	**Odds Ratio (95% Confidence Intervals)**	**P value**
Age (Years)	**1.04** (1.01 - 1.06)	0.001	**1.02** (0.98 - 1.06)	0.24
Menopausal status (pre vs post)	**4.49** (2.49 - 8.09)	0.001	**1.92** (0.66 – 5.61)	0.23
Ultrasonographic Endometrial thickness (5mm cut-off)	**0.15** (0.06 – 0.37)	0.001	**0.19** (0.07 – 0.53)	0.001
Hysteroscopic endometrial thickness (atrophic vs non-atrophic)	**5.53** (2.47 – 11.60)	0.001	**4.79** (1.05 - 21.91)	0.04

*Multivariable logistic regression model with insufficient or sufficient sample on endometrial biopsy as the binary dependent variable and age, menopausal status, atrophic endometrium on hysteroscopy, and endometrial thickness on sonography as the independent variables. See methods for details. Parity (odds ratio 1.2, 95% confidence interval 0.9-1.5, β-coefficient 0.04, p=0.8) and use of hormone replacement therapy (odds ratio 1.6, 95% confidence interval 0.8-2.8, β-coefficient –0.24, p=0.9) did not contribute to the model. There was no forward or backward selection or rejection process for the variables, they were forced into the analysis on clinical relevance.
(Model deviance=98.46; Hosmer-Lemeshow goodness of fit test p=0.43 at 8DF)

The Hosmer and Lemeshow Goodness-of-Fit Test below divides subjects into deciles based on predicted probabilities, then computes a chi-square from observed and expected frequencies. The p-value=0.43 is computed from the chi-square distribution with 8 degrees of freedom (8DF) and indicates that the logistic model is a good fit. That is, if the Hosmer and Lemeshow Goodness-of-Fit test statistic is .05 or less, we reject the null hypothesis that there is no difference between the observed and predicted values of the dependent; if it is greater, as we want, we fail to reject the null hypothesis that there is no difference, implying that the model's estimates fit the data at an acceptable level.

4.4 Discussion

4.4.1 Principal findings of the study:

This study showed that insufficient sample on endometrial biopsy was associated with hysteroscopic finding of endometrial atrophy and ultrasonographically thin endometrium. Hysteroscopic finding of atrophy is subjective whereas measurement of endometrial thickness is objective. However, both of these tests independently influenced the odds of insufficient sample on outpatient biopsy making it more likely when there was evidence of endometrial atrophy.
There was no association with the patient's age, menopausal status, parity and use of hormone replacement therapy in multivariable analysis.

4.4.2 Appraisal of quality of the study

The type of controlled analysis used in this Chapter to delineate the significance of factors associated with insufficient specimen on outpatient endometrial biopsy has not been reported so far.

The strength of this study and the credibility of its findings depend on the strength of the methodology. This is made clear by applying the Standard of Quality for Reporting Logistic Regression Models (Khan 1999) the objective of which is to to identify problems affecting the accuracy and interpretation of this approach to multivariable statistical analyses. Table 4.3.

Table 4.3: Application of the Standard of Quality for Reporting Logistic Regression Models

Quality Feature	Description of the Criteria	Compliance of This Study
Did the model appear to be correct?		
• Research hypothesis	It should be reported clearly	Yes
• Reporting of the variable selection process	It is adequate if it is based on consideration of the clinical importance of various factors	Yes
• Statistical significance threshold	The threshold of statistical significance for inclusion or deletion of variables (e.g. in stepwise forward or backward regression) should be reported	Yes
• Conformity to linear gradient	An attempt should be made to detect this problem, e.g. by comparison of the observed and the predicted values for the outcome over the ordinal zones or by an alternative analysis using cross-stratification	Yes
How well did the overall model work?		
• Testing for goodness-of-fit	The accuracy with which the regression model describes the data should be assessed	Yes
How important was each of the independent variable?		
• Unit of measurement	The change in dependent variable associated with a unit change in the independent variable should be reported	Yes
• Coding of variables	A description of variables if they were on an interval, ordinal or binary scale should be provided	Yes
• Testing for interactions	when the impact of one variable is dependent on the level of another variable, interactions between independent variables should be examined	Yes

4.4.3 Relation to other studies considering differences in quality and results

Generally, an insufficient sample in women with abnormal uterine bleeding is not believed to be a cause for concern as failure to obtain an endometrial specimen from a devise correctly positioned within the uterine cavity is considered to be an assurance that no significant intrauterine pathology is present. There is still however controversy about the need for additional testing if outpatient endometrial biopsy is insufficient because of the lack of confidence that insufficient sample is benign. Some gynaecologists prefer to perform inpatient curettage under anaesthesia while others suggest that further tests may not be necessary (Ben Baruch *et al.*, 1994).

4.4.4 Practical implications for clinicians

It would appear from this analysis that clinicians can be confident in reassuring women with insufficient sample on outpatient endometrial biopsy provided the hysteroscopic and sonographic endometrial assessment is consistent with endometrial atrophy. This means that inpatient dilatation and curettage can possibly be avoided in such women. However, in majority of settings the clinicians will only have the ultrasound scan result and perhaps not the outpatient hysteroscopy, in such a situation, the same advice still prove correct (i.e. to reassure and discharge women with insufficient endometial sample with thin endometrium on scan) without the need to book then for inpatient hysteroscopy and curettage. This is because p-value for the ultrasonographic endometrial thickness did not change on performing the multivariable regression analysis (p=0.001), as much as it did with hysteroscopic endometrial thickness (from p=0.001 to p=0.04). Table 4.2. Keeping in mind the need for usual precautions of reinvestigation on recurrence of symptoms.

4.5.4 Implications on research

- Insufficient sample may be regarded as a negative result in metaanalysis when outpatient biopsy is used as a gold standard.

- There is a suggestion that failure to obtain a sufficient endometrial sample in an outpatient biopsy may be due to the improper positioning of the Pipelle sampler devise within the uterine cavity. Therefore there is a need to design a randomised controlled trial looking at the success rate of obtaining sufficient sample with or without the use of ultrasound scan at the time of sampling to monitor for the correct positioning of the Pipelle sampler devise.

CHAPTER 5

THE RISK OF PREMALIGNANT AND MALIGNANT PATHOLOGY IN ENDOMETRIAL POLYPS

5.1 Introduction

Abnormal uterine bleeding, accounting for up to 20% of outpatient visits, is one of the most common gynaecological problems (Ben Yehuda *et al.,* 1998). In women over the age of 40 years, the management requires prompt diagnosis to exclude intrauterine pathology. Traditionally, dilatation and curettage was performed to exclude endometrial pathology. However, as this is a blind procedure, endometrial polyps have been missed in 50-85% of cases (Maja 1996; Maja Jr 1996). Currently outpatient evaluation is preferred using imaging techniques including hysteroscopy and ultrasonography (Bakour 1999; Cicinelli 1993; Gimpleson 1998; Towbin 1996). These new methods of evaluation have revealed that uterine polyps are common findings, ranging in rate from 16 to 34% (Gimpleson RJ 1998; Loffer 1989; Saidi, Sadler, Theis, Akright, Farhart, Villanueva 1997; Towbin NA 1996; Valle 1981). The aetiology and pathogenesis of endometrial polyps are not fully understood but polyps are believed to be a risk factor for endometrial cancer (Maja Jr H 1996). However, this issue has seldom been substantiated in a controlled study. With the knowledge that only a few studies have looked at the pathological significance of endometrial polyps (Maja 1996; Maja Jr 1996), this study was conducted to comprehensively address the following research question:

What is the risk of premalignant and malignant pathology among endometrial polyps?

5.2 Methods

A cohort of 248 consecutive patients was prospectively evaluated for abnormal uterine bleeding in a rapid access One-stop outpatient clinic from November 1996 to December 1997 (Bakour 1999). Ethical considerations are discussed in Section 2.2.

The One-stop clinic was conducted by a consultant or by a senior minimal access surgery trainee. After consultation, counselling and physical examination, outpatient mini-hysteroscopy, and Pipelle endometrial sampling were performed. Inpatient hysteroscopy and curettage under general anaesthesia were performed if the hysteroscopic view was poor or if the biopsy was not obtained in outpatient setting. Pelvic transvaginal ultrasonography was performed in some patients using an endovaginal 6.5 MHz transducer of a portable scanner (Hitachi Sumi, Japan) (Bakour 1999). Outpatient hysteroscopy was performed using a 1.2mm micro-hysteroscope in 2.5mm outer rigid sheath (Karl Storz, Tuttlingen, Germany) with normal saline solution as the distending solution. Outpatient endometrial biopsies were performed using Pipelle sampling (Laboratoire CCD, Paris, France) (Stovall 1991). The polyps were all removed under general anaesthesia by avulsion with a polypectomy forceps and subjected to histological assessment. Histological diagnosis was classified as benign (hormonal, exogenous hormonal effect, simple polyp), hyperplasia (simple, complex or atypical) or carcinoma. Some reviewers would not agree to include simple hyperplasia cases in the analysis on the categorisation that simple hyperplasia is not a premalignant lesion. However, in my opinion, simple hyperplasia is still a pathological histology and clinically if symptomatic would need treatment.

To determine the magnitude of malignant potential among polyps, the pathological findings in polyps (cases) were compared with non-polypoidal specimens (controls) using Fisher's Exact Test. The risk of malignancy among polyps with estrogenic exposure was also examined.

5.3 Results

Of the 248 patients, 62 (25%) had endometrial polyps on hysteroscopy. The mean age was 53.7 (95% CI 49.6-57.8), the mean parity was 2.2 (95% CI 1.8-2.6). Of the 62 women with polyps, 30 (48.4%) were postmenopausal, 19 (30.6%) women were on hormone replacement therapy (HRT) and 10 (16.1%) were receiving tamoxifen. Histologically, 53/62 (85.5%) polyps were benign, 6/62 (9.6%) had hyperplasia (two simples and four complex/atypical), and 3/62 (4.8%) polyps were malignant.

Overall 24/248 patients had hyperplasia or cancer on histology in this series. Of these, nine cases were found in polyps and 15 were found in women without polyps (9/62 vs 15/186, 14.5% vs 8.1%). As shown in Table 5.1, hyperplasia was more frequent among women with polyps than in women without (9.6% vs 4.8%, OR= 2.1, 90% CI 0.85-5.2). This result was not significant at the traditional cut-off p-value as the p value was 0.16. However, it shows a significant trend towards an association between hyperplasia and polyps. There was no such difference for cancer (3.4% vs 2.8%, OR= 1.5, 90% CI 0.46-5.0, p=0.5).

Table 5.1: Endometrial hyperplasia and cancer in women with endometrial polyps and those without

Histology*	Without Polyp N=186	With Polyps N=62	P Value#	Odd Ratio 90% CI
Hyperplasia	9 (4.8%)	6 (9.7%)	0.16	2.1
-Simple	3 (1.6%)	2 (3.2%)		(0.85-5.2)
-Complex/Atypical	6 (3.2%)	4 (6.5%)		
Carcinoma	6 (3.2%)	3 (4.8%)	0.5	1.5 (0.46-5.0)

*Histologic diagnoses are classified as benign (including hormonal effect), hyperplasia (simple, complex, or atypical) or cancer based on the pathology report.
P value for Chi- square Test comparing the effect of polyps on the occurrence of hyperplasia or cancer on histopathology.

Separate analysis (not included in this thesis) carried out among the same cohort of the 248 patients, showed a high incidence of polyps in those taking tamoxifen (tamoxifen group) with approximately 36% of these polyps showing abnormal changes. This background prompted the examination of the relationship between estrogenic exposure and risk of malignancy among polyps. This further analysis revealed that amongst the 19 HRT users one patient had complex hyperplasia and one had atypical hyperplasia, while amongst the 10 tamoxifen users two had complex/atypical hyperplasia. No cancer cases were found (Table 5.2). Compared to no estrogenic exposure, HRT and tamoxifen exposure increased the odds of having polyps with hyperplasia or cancer but this result was not statistically significant (OR=1.5, 95% CI 0.36-6.26, p=0.5).

Table 5.2: Histology of endometrial polyps and estrogenic exposure to hormone replacement therapy or tamoxifen

Histology*Of Endometrial Polyps	Estrogenic exposure		No exposure	P value#
	HRT+	Tamoxifen		
Benign	17	7	29	
Hyperplasia				
- Simple	0	1	1	0.5#
-Complex/atypical	2	2	0	
Cancer	0	0	3	

* Histologic diagnoses are classified as benign (including hormonal effect), hyperplasia (simple, complex, or atypical) or cancer based on the pathology report.

+ HRT = hormone replacement therapy

P value for Fisher Exact Test comparing the effect of estrogenic exposure (HRT and tamoxifen) vs non-exposure on the occurrence of benign vs hyperplasia/cancer. (OR=1.5, 95% CI 0.36-6.26, p=0.5).

5.4 Discussion

5.4.1 Principal findings of the study

The principal finding of this study was that in polyps the rate of abnormal histology was approximately double that in the non-polyp group. The study concluded that hyperplasia was, but cancer was not, more common in women with endometrial polyps than in those without.

5.4.2 Appraisal of quality of the study

The credibility of these findings depends on the strength of the methodology. Polyps were identified using high quality imaging with microfibreoptic hysteroscopes. Hysteroscopy was considered to be the gold standard for detecting endometrial polyps because it provided direct visualisation of the endometrial cavity and polyps in it (Bakour 1999; Haller 1996; Loffer 1989; Saidi 1997; Valle 1981). Hence, any chance of misclassification of risk category was limited to a minimum. As far as the outcome variable of the index study is concerned, the nature of pathology was confirmed in all patients by histological examination of specimens obtained from the endometrial cavity. As histology was the gold standard for diagnosing hyperplasia or cancer, the risk of misclassification of outcome was also eliminated. In this manner, the methodology was robust to produce a credible estimate of the malignant potential associated with endometrial polyps.

The association of endometrial polyps with premalignant and malignant pathology has not been presented before in a control analysis. Its conclusion may, however, be challenged on the basis of variation in p-values (p value= 0.16, OR= 2.1). This challenge would be strongly rebutted on the basis that the significance level used in hypothesis testing is based on an arbitrary dichotomy and it is not as informative as the CI (Brennan 1994). In fact, Sterne *et al.,* (2001) have recommended that p values should not be used in observational studies and 90% confidence intervals around odds ratios should be readily accepted. However, the results showed a significant trend towards an

association between hyperplasia and polyps. Furthermore, the true values are more likely to lie near the centre of the interval than at the extremes i.e. the true point estimate is more likely to be >1.0.

5.4.3 Relation to other studies considering differences in quality and results

It appears that the evidence is suggestive of a clinically significant relationship between polyps and hyperplasia that merits consideration. This trend has been confirmed in an independent data set collected in the same One -Stop clinic subsequent to this paper involving 790 women. Hyperplasia was found more frequent in women with polyps than in those without (14/95 vs 38/695 or 14.73% vs 5.46%, p=0.0006, OR=2.98, 90% CI 1.72-5.17) while there was no significant difference with regard cancer (1/95 vs 19/695 or 1.05% vs 2.73%, p=0.33, OR=0.37, 90% CI 0.07-2.06) (personal communication).

Moreover, on examining the consistency of the findings of this study with that of other literature, there was satisfaction that the 25% rate of endometrial polyps was comparable with the 16-34% range of rates quoted by other authors (Gimpleson RJ 1998; Loffer 1989; Saidi, Sadler, Theis, Akright, Farhart, Villanueva 1997; Towbin NA 1996; Valle 1981). The increasing use of HRT in postmenopausal women and adjuvant therapy (tamoxifen) in women with breast cancer (Maja 1996; Maja Jr 1996) was believed to be associated with endometrial abnormalities. HRT may induce polyps or it may simply stimulate pre-existing polyps due to the effect of estrogens. Tamoxifen administration as adjuvant therapy for breast cancer, confers a survival benefit (Cheng 1997; Exacoustos 1995; Ismail 1996; Maja 1996; Maja Jr 1996). However, the drug's weak estrogen effect may lead to the development of endometrial proliferative abnormalities ranging from hyperplasia to invasive cancer, with endometrial polyps occupying the intermediate ground (Cheng 1997; Cohen 1993; Hardell 1988; Lahti 1993). This background prompted the examination of the relationship between estrogenic exposure and the risk of malignancy among polyps. It was found

that the odds of hyperplasia and cancer were high with estrogenic exposure (Table 5.2). Although the result from Table 5.2 did not reach statistical significance, it supports the policy that symptomatic women on HRT and tamoxifen should be investigated promptly (Cheng 1997; Cohen 1993; Hardell 1988; Lahti 1993; Powles 1998)

5.4.4 Practical implications for clinicians

The diagnosis of uterine polyps is becoming more common than previously assumed. This study indicates that the detection of polyps should be taken seriously. The management of polyps is at present varied but the findings of this study reinforce the need for their formal removal followed by histological assessment.

5.4.5 Implications on research

- This study revealed that in abnormal uterine bleeding, hyperplasia was more common in women with endometrial polyps than in those without. Therefore the diagnosis, prognosis and management of polyps require further evaluation in larger study in order to avoid the risk of type-II error (where there is in fact a difference between the two groups but the study failed to reveal it as it was not powered enough).
- Randomised trials are required to address the effectiveness and cost-effectiveness of various methods of polypectomy.

CHAPTER 6

RISK FACTORS ASSOCIATED WITH ENDOMETRIAL POLYPS IN ABNORMAL UTERINE BLEEDING

6.1 Introduction

Endometrial polyps are a common problem in women with abnormal uterine bleeding accounting for up to 25% of such cases (Bakour *et al.,* 2000). Hormone replacement therapy (Maja Jr 1996) (with or without progestogen) and tamoxifen treatment (Cheng 1997) have been associated with the development of endometrial polyps, presumably mediated through agonistic estrogenic effects (Reslova *et al.,* 1999). However, polyps are also frequently seen in postmenopausal women who are not exposed to these estrogenic agents. In the relationship between exposure to estrogenic agents and occurrence of polyps, the confounding effects of other factors have not been taken into account. Therefore, a controlled analysis was conducted to answer the following research question:

What is the significance of various risk factors (age, parity, menopausal status, hormone replacement therapy, and tamoxifen) in the development of endometrial polyps?

By evaluating the relationship of hormone replacement therapy and tamoxifen treatment with the development of endometrial polyps, adjusting for the confounding effects of other factors such as age, parity and menopausal status.

6.2 Methods

To perform this analysis I used a database of 248 consecutive patients evaluated for abnormal uterine bleeding in the One-stop from November 1996 to December 1997 (Bakour 1999). Ethical considerations were discussed in Section 2.2.

Patients were referred to the service from gynaecology outpatient clinics or directly from general practice for the investigation of a wide spectrum of abnormal uterine bleeding including: menorrhagia, irregular menstrual loss, postmenopausal bleeding, unscheduled bleeding on tamoxifen, and unscheduled bleeding on combined hormone replacement therapy either sequential or continuous.

Endometrial polyps were diagnosed using outpatient hysteroscopy performed with a 1.2 mm micro-hysteroscope in a 2.5mm rigid sheath (Karl Storz, Tuttlingen, Germany) with normal saline solution as the distension medium. Direct visualization using hysteroscopy is generally regarded as the 'gold standard' for the diagnosis of polyps (Loffer 1989; Pal *et al.,* 1997). Submucous fibroid polyps and thickened polypoidal endometrium were not considered as endometrial polyps. For statistical analysis, each patient was regarded as a single unit of analysis whether with single, multiple, small or large polyps. The histological nature of the polyp was confirmed in all patients by performing inpatient polypectomy.

To explore whether tamoxifen or hormone replacement therapy were associated with polyps while adjusting for the effects of age, parity and menopausal status, a multivariable logistic regression model was built with the occurrence of polyps as the binary outcome variable. The primary null hypothesis was, in terms of the logistic model "the value of the beta coefficient for the exposure to estrogenic therapy is zero or its exponent is 1.0". To obtain a valid estimate of the association between exposure to estrogenic therapy (hormone replacement therapy and tamoxifen) and occurrence of polyps, the analysis was controlled for the confounding effects of age, parity and menopausal status by including them in the multivariable logistic regression analysis. The model building process was undertaken according to published guidelines (Khan 1999). The model that

best predicted the occurrence of polyps with the lowest deviance and highest goodness of fit was selected.

6.3 Results

Sixty-two cases with endometrial polyps were identified by hysteroscopy. The characteristics of the women attending the One-stop clinic are shown in Table 6.1. This table provides a simple comparison between women with polyps (cases) and those without polyps (controls) for age, parity, menopausal status, tamoxifen use and hormone replacement therapy.

Table 6.1: Comparison of women with and without polyps diagnosed in One-stop clinic

Patients' Characteristics	With Polyp N=62	Without Polyp N=186
Age - years Mean (95% CI)	53.7 (49.6-57.8)	49.6 (47.5-51.8)
Parity Mean (95% CI)	2.2 (1.8-2.6)	2.5 (2.3-2.7)
Menopausal Status		
Postmenopausal	31 (50%)	81 (43.5%)
Premenopausal n (%)	31 (50%)	105 (56.5%)
Use of Hormone Replacement Therapy		
Yes	19 (30.6%)	54 (29%)
No n (%)	43 (69.4%)	132 (71%)
Tamoxifen Treatment		
Yes	10 (16.1%)	5 (2.7%)
No n (%)	52 (83.9%)	181 (97.3%)

Univariable regression (unadjusted analysis) revealed that in this group of women undergoing investigation for abnormal uterine bleeding, tamoxifen treatment increased the odds of having an endometrial polyp (OR=6.96, p=0.0007). The effect of hormone replacement therapy and menopausal status did not statistically significantly increase the odds of having a polyp (OR=1.08, p=0.80 and 1.25, p=0.44 respectively). Multivariable regression analysis adjusting for the confounding effects of age, parity and menopausal status showed that the effect of tamoxifen treatment on occurrence of polyps continued to be significant (adjusted OR 11.21, p=0.0009), but the use of hormone replacement therapy did not have a significant association (adjusted OR=1.48, p=0.32). Table 6.2.

Table 6.2: Logistic regression analysis to determine the significance of factors associated with endometrial polyps.

Risk Factor	Univariable Regression		Multivariable Regression*	
	Odds Ratio (95% Confidence Intervals)	**P-value**	**Odds Ratio (95% Confidence Intervals)**	**P-value**
Age (1 year increment)	**1.02** (1.00 - 1.04)	0.07	**1.00** (0.98 - 1.03)	0.80
Parity (1 unit increment)	**0.86** (0.69 - 1.08)	0.20	**0.83** (0.64 - 1.07)	0.15
Menopause (Post / Pre)	**1.25** (0.70 – 2.24)	0.44	**0.95** (0.44 - 2.04)	0.89
Hormone Replacement Therapy (Yes / No)	**1.08** (0.58 - 2.02)	0.80	**1.48** (0.68 - 3.20)	0.32
Tamoxifen Therapy (Yes / No)	**6.96** (2.28 - 21.26)	0.0007	**11.21** (2.70 – 46.46)	0.0009

*Multivariable logistic regression model with presence or absence of polyp as the binary outcome variable and age, parity, menopausal status, hormone replacement therapy and tamoxifen treatment as predictor variables. See methods for details. There was no forward or backward selection or rejection process for the variables, they were forced into the analysis on clinical relevance.
(Model deviance=98.46; Hosmer-Lemeshow goodness of fit test p=0.43 at 8DF)

The Hosmer and Lemeshow Goodness-of-Fit Test below divides subjects into deciles based on predicted probabilities, then computes a chi-square from observed and expected frequencies. The p-value=0.43 is computed from the chi-square distribution with 8 degrees of freedom (8DF) and indicates that the logistic model is a good fit. That is, if the Hosmer and Lemeshow Goodness-of-Fit test statistic is .05 or less, we reject the null hypothesis that there is no difference between the observed and predicted values of the dependent; if it is greater, as we want, we fail to reject the null hypothesis that there is no difference, implying that the model's estimates fit the data at an acceptable level.

6.4 Discussion

6.4.1 The principle findings of this study

This study confirmed that tamoxifen was associated with endometrial polyps. However, it rejected the hypothesis that hormone replacement therapy was a risk factor for endometrial polyps.

6.4.2 Appraisal of quality of the study

The strength of this study and the credibility of its findings depend on the strength of the methodology. This is made clear by applying the Standard of Quality for Reporting Logistic Regression Models (Khan 1999) the objective of which is to to identify problems affecting the accuracy and interpretation of this approach to multivariable statistical analyses. Table 6.3.

Table 6.3: Application of the Standard of Quality for Reporting Logistic Regression Models

Quality Feature	Description of the Criteria	Compliance of This Study
Did the model appear to be correct?		
• Research hypothesis	It should be reported clearly	Yes
• Reporting of the variable selection process	It is adequate if it is based on consideration of the clinical importance of various factors	Yes
• Statistical significance threshold	The threshold of statistical significance for inclusion or deletion of variables (e.g. in stepwise forward or backward regression) should be reported	Yes
• Conformity to linear gradient	An attempt should be made to detect this problem, e.g. by comparison of the observed and the predicted values for the outcome over the ordinal zones or by an alternative analysis using cross-stratification	Yes
How well did the overall model work?		
• Testing for goodness-of-fit	The accuracy with which the regression model describes the data should be assessed	Yes
How important was each of the independent variable?		
• Unit of measurement	The change in dependent variable associated with a unit change in the independent variable should be reported	Yes
• Coding of variables	A description of variables if they were on an interval, ordinal or binary scale should be provided	Yes
• Testing for interactions	when the impact of one variable is dependent on the level of another variable, interactions between independent variables should be examined	Yes

6.4.3 Relation to other studies considering differences in quality and results

While the aetiology and pathogenesis of endometrial polyps are not fully understood, it is generally believed that hormone replacement therapy (with or without progestogen) is a factor that may induce endometrial polyps (Maja Jr 1996). The results of this study using robust analysis do not support this notion.

When evaluating the association between hormone replacement therapy use and occurrence of polyps, multivariable logistic regression analysis was used in this study to adjust for the confounding effect of several variables. For example, inclusion of menopausal status in the model produced a stratified analysis of the effect of hormone replacement therapy on polyps in premenopausal and postmenopausal women, and included an adjustment for this confounding variable. In this way, the odd ratios (and their p-values) demonstrated the lack of an association between hormone replacement therapy and polyps in light of the variation in menopausal status, age, parity, and tamoxifen use in this sample. In addition, these results confirmed that tamoxifen treatment was a significant risk factor associated with endometrial polyps.

6.4.4 Practical implications for clinicians

There are numerous recognized reasons for long-term non-compliance use of hormone replacement therapy (Coope 1992), such as irregular bleeding and/or fear of increased risk of breast cancer. It appeared from this study that the risk of hormone replacement therapy inducing endometrial polyps should not be one of these reasons.

Clinicians have to be careful in interpreting these findings since the study population was a selected group of women who were referred to the special clinic for the investigation of abnormal uterine bleeding. This means that the generalization of the results is limited to symptomatic women.

6.4.5 Implications on research

It appeared from this study that the risk of hormone replacement therapy inducing endometrial polyps should not be one of the reasons for long-term non-compliance use of hormone replacement therapy. It is important to confirm the lack of the association between hormone replacement therapy and polyps in a larger study to inform the debate for and against the use of hormone replacement therapy.

CHAPTER 7

FEASIBILITY OF MULTIVARIABLE ANALYSIS TO EVALUATE COMBINATIONS OF TESTS IN THE DIAGNOSIS OF ENDOMETRIAL DISEASE

7.1 INTRODUCTION

Accurate diagnosis is a critical part of the clinical process because it allows optimal management strategies to be employed. Incorrect diagnoses may put patients at risk and waste limited resources (Fineberg 1979; Reid 1995). Medical diagnosis is usually based on information acquired from a variety of components, which make up the clinical process (i.e. history, examination and tests). The need to rigorously assess the accuracy of the clinical history and examination has been highlighted (McAlister 1999), but in our view the need is to integrate this part of the clinical process with diagnostic tests. The whole clinical process should be borne in mind when conducting research to evaluate a diagnostic strategy and when implementing new tests into clinical practice. Diagnostic accuracy in studies evaluating tests in isolation from the rest of the clinical context does not necessarily indicate how useful the test will be in practice (Begg 1989; Khan 2001; Moons ***et al.,*** 1996). To determine this, evaluation within the context of contemporary clinical practice is essential. This is because the information generated by a diagnostic test may have already been obtained from the patient's history and physical examination, which have taken place earlier in the clinical process. The true clinical value of a test lies in the added information over and above what was already known from the history and examination.

A stepwise multivariable approach in the evaluation of diagnostic tests (Khan 1999) allows the

clinical context, in which the test will be used, to be accounted for. This study develops such an approach and explores its application in the diagnosis of endometrial disease (hyperplasia or cancer) in the context of an outpatient rapid access clinic. In the same unit systematic reviews (Clark 2002a; Clark 2002b; Clark *et al.,* 2001; Gupta 2002) were conducted addressing the diagnostic accuracy of tests used in predicting endometrial disease in this setting (ultrasound, hysteroscopy and biopsy), which included over 100 primary studies, none of which used stepwise multivariable approaches. Therefore, the research question raised in this chapter was:

What is the feasibility of multivariable analysis to evaluate the added value of combinations of diagnostic tests in the diagnosis of endometrial disease in routine practice where some diagnostic information is already available from clinical history?

7.2 Methods

In an outpatient One-stop clinic for investigation of abnormal uterine bleeding, patients' age at presentation, menopausal status and use of hormone replacement therapy are recorded. The ethical considerations are covered in Section 2.2. Patients then undergo further investigation with a combination of pelvic ultrasound scan and outpatient hysteroscopy. Data collected from 248 consecutive patients investigated in our one-stop clinic for abnormal uterine bleeding from November 1996 to December 1997 was used to develop a model to determine the added value of ultrasonography and hysteroscopy.

Univariable analyses of the tests employed in this data set have been published Previously (Bakour 1999a; Bakour 1999b). Pelvic transvaginal ultrasonography was performed using an endovaginal 6.5 MHz transducer of a portable scanner (Hitachi Sumi, Japan). Double layer endometrial thickness was measured in mm, and endometrial thickness of 5mm was used as a cut-off, based on the findings of a recent review (Smith-Bindman *et al.,* 1998). Mini-hysteroscopy was performed

using a 1.2mm micro-hysteroscope with 2.5mm rigid sheath (Karl Storz, Tuttlingen, Germany). By applying simple pressure with the hysteroscope, the depth of mucosa was assessed; hysteroscopic features of smooth non-vascular endometrium were considered as normal hysteroscopy while features of increased endometrial thickness, abnormal vascularisation, irregular friable polypoid formations with necrosis or bleeding, were considered suspicious lesions (Ben Yehuda *et al.,*1998; Cicinelli 1993; Uno *et al.,* 1995). The gold standard was based on histological diagnoses. These diagnoses were classified as either *negative* (secretory/proliferative endometrium and benign endometrial polyps), or *positive* (endometrial hyperplasia and carcinoma), because the main aim of investigations for abnormal uterine bleeding is to exclude serious intrauterine pathology, primarily endometrial cancer and hyperplasia. The secondary aim is to stop the bleeding which is the real symptomatic problem to the patient.

Histological samples were provided by outpatient endometrial biopsy (Laboratoire CCD, Paris, France), dilatation and curettage (D&C) and hysterectomy specimens. In patients where outpatient biopsy was inadequate, inpatient sampling was performed. All patients were followed up for a minimum of 6 months and recurrent symptoms were used as an indication for reassessment. Although outpatient biopsy is a simple and inexpensive diagnostic test, there are concerns surrounding the non-representative nature of blind endometrial biopsy (Guido 1995; Rodriguez 1993) and this is the rationale for employing hysteroscopic or ultrasound imaging. However, in this study, no erroneous diagnoses on outpatient biopsy were identified from hysterectomy specimens, D&C or following reinvestigation of symptoms at 6 months follow up and so its use as the diagnostic reference standard appears to be justified.

Logistic regression analysis (Khan 1999; Norman 1994) was used to delineate the predictive values of the historical features age, menopausal status and hormone replacement therapy use, and results

from ultrasound and hysteroscopy (independent variables) on endometrial hyperplasia or cancer (binary dependent variable).

Initially univariable analysis was performed followed by multivariable modeling. Four diagnostic models were produced by multivariable analysis, which included historical features alone, historical features plus ultrasonography, historical features plus hysteroscopy and historical features plus ultrasonography and hysteroscopy. Postmenopausal status, use of hormone replacement therapy, ultrasonic endometrial thickness greater than 5mm and suspicious findings on hysteroscopy were all considered positive results. Age was split into six categories; less than or equal to 30 years, 31-40 years, 41-50 years, 51-60 years, 61-70 years and over 70 years. Increasing age was considered more likely to be associated with endometrial pathology.

The objective of this study was to develop an analytical approach. To allow the building of multivariable models using this small data set, cases of endometrial cancer and hyperplasia had to be used combined to increase the number of outcome events and any missing data were imputed (Norman 1994) (see results). Using a univariable approach for each diagnostic (independent) variable, the diagnostic odds ratio (dOR) was calculated where values >1.0 showed an increased level of accuracy. The dOR represents a ratio of the positive and negative likelihood ratios and it can be mathematically summarized as:

dOR = [sensitivity /(1-specificity)] / [specificity /(1-sensitivity)].

After the univariable analysis, four multivariable logistic regression models were then built (Concato 1993; Khan 1999; Norman 1994) using BMDP statistical software release (Concato 1993; Norman 1994) (BMDP Statistical Software Inc., Los Angeles, CA). These models were built to allow the statistical approach to mimic the clinical process. The first model built was to provide a valid estimate of the combined predictive value of the historical variables (i.e. age, menopausal

state and use of hormone replacement therapy). The other three models determined the combined predictive added value of history and tests (ultrasonography or hysteroscopy and ultrasound or hysteroscopy combined). This mirrors the real clinical situation where historical information is acquired prior to undertaking investigation with ultrasound scan or hysteroscopy.

The predictive value of each model was summarized by plotting the estimates of sensitivity (true positive rates) against 1- specificity (false positive rates) to develop a receiver operating characteristic (ROC) curve that characterizes the performance of the test (Hanley 1982; Swets 1988). This is a commonly used approach in both primary and secondary research for evaluating diagnostic accuracy of tests (Irwig 1995; Linnet 1988; Swets 1988; Walter 1999). The ROC curves can be readily generated from logistic regression modelling and this type of approach has been used in tests with multiple covariates (Tosteson 1994). The area under the curve (ROC area) determines the accuracy with which the test diagnoses the condition of interest. A ROC area greater than 0.5 suggests some degree of test accuracy, with higher accuracy suggested by an ROC area closer to 1.0 (representing perfect test accuracy). The ROC areas for each model were compared.

The hypothesis that the study designed to test was about the improvement in prediction achieved by adding tests to the history. The chi-square test was used to test for statistically significant improvement between models and was computed from the log of the ratio of the current versus the previous likelihood function values. A small p-value indicated a significant change in prediction.

7.3 Results

The mean age of the 248 women included in this study was 50 years (range 23 to 94 years) and 112 (45%) were postmenopausal. Forty- five women (18%) were under 40 years of age and 73 women (30%) were taking hormone replacement therapy. Twenty-three women (9%) had endometrial

disease (cancer/hyperplasia) on histopathological assessment. All cases of cancer detected were in women over 40 years of age as were the majority of endometrial hyperplasia (85%). Data for hysteroscopy and outpatient biopsy were complete, but 10% of the total data were missing for the other variables and these data were imputed (Norman GR 1994). There were 14 false-positive and 17 false-negative diagnoses with hysteroscopy and 121 false-positive and seven false-negative diagnoses with ultrasound measurement of endometrial thickness.

Table 7.1 shows the results of the univariable and multivariable approaches to summarize the diagnostic value of history and tests. The historical features (age, postmenopausal bleeding and use of hormone replacement therapy) were all significant predictors of endometrial hyperplasia or cancer using a univariable approach. Similarly, ultrasonography and hysteroscopy were also significant predictors. These univariable evaluations were then combined in a stepwise fashion in keeping with the clinical context described above.

The predictive ability of all the historical features combined measured by the ROC area for the model including historical features was 0.78. The addition of ultrasonography to historical features significantly increased the ROC area (0.78 versus 0.82, χ^2 for improvement = 5.6, p=0.02) (Figure 7.1). Similarly, the addition of hysteroscopy to historical features significantly increased the ROC area (0.78 versus 0.81, χ^2 for improvement = 7.1, p=0.008) (Figure 7.2). The addition of both ultrasonography and hysteroscopy increased the ROC area to 0.84. This represented significantly improved predictive ability from the clinical history plus ultrasonography model (0.82 versus 0.84, χ^2 for improvement = 6.9, p=0.009) and from the clinical history + hysteroscopy model (0.81 versus 0.84, χ^2 for improvement = 5.4, p=0.02) (Figures 7.1 and 7.2).

Table 7.1: Results of univariable and stepwise multivariable analyses to assess the ability of patient history, hysteroscopy and ultrasonography to predict the presence of endometrial hyperplasia or cancer in 248 women with abnormal uterine bleeding.

Diagnostic model	**Univariable analysis**		**Multivariable analysis**											
			Clinical history			**Clinical history + ultrasound**			**Clinical history + hysteroscopy**			**Clinical history + ultrasound + hysteroscopy**		
Predictor variables (cut-off levels)	Unadjusted dOR (95% CI)	P value	Adjusted dOR (95% CI)	P value	ROC area	Adjusted dOR (95% CI)	P value	ROC area	Adjusted dOR (95% CI)	P value	ROC area	Adjusted dOR (95% CI)	P value	ROC area
Clinical history														
***Age:**	-		-			-			-			-		
(31 – 40 years	0.7 (0.04-12.3)		1.0 (0.05-18.0)			0.7 (0.04-13.2)			0.9 (0.05-17.5)			0.6 (0.03-12.0)		
41 50 years	0.9 (0.1-9.1)		1.1 (0.1-10.9)			1.1 (0.1-10.8)			1.0 (0.1-10.5)			0.9 (0.09-9.4)		
51 – 60 years	0.6 (0.06-5.9)	0.005	0.5 (0.04-5.2)	0.008		0.3 (0.03-4.0)	0.06		0.4 (0.04-5.1)	0.15		0.3 (0.02-3.44)	0.09	
61 – 70 years	4.9 (0.06-42.7)		3.0 (0.3-32.9)		**0.78**	2.2 (0.2-26.3)			2.3 (0.2-26.8)			1.5 (0.19-19.1)		
> 70 years)	5.1 (0.5-52.9)		2.3 (0.2-30.3)			2.1 (0.2-30.8)		**0.82**[2]	2.4 (0.2-33.0)		**0.81**[1]	2.0 (0.13-30.0)		**0.84**[3]
Menopausal status (pre vs post)	4.0 (1.5-10.5)	0.003	3.4 (0.8-14.2)	0.08		4.2 (1.0-19.6)	0.05		3.4 (0.8-14.4)	0.09		4.4 (1.0-20.7)	0.05	
Hormone replacement therapy (use vs non-use)	0.3 (0.09-1.0)	0.05	0.3 (0.07-0.99)	0.04		0.3 (0.07-0.99)	0.03		0.3 (0.08-1.2)	0.06		0.3 (0.08-1.2)	0.07	
Additional diagnostic procedures														
Ultrasound endometrial thickness (5mm cut-off)	2.61 (1.0-6.9)	0.04	0.61	-		3.4 (1.2-9.8)	0.02					5.3 (1.6-16.7)	0.009	
Hysteroscopy (suspicious vs normal)	7.1 (2.5-20.3)	0.001	0.62	-					5.3 (1.6-17.0)	0.008		3.5 (1.1-10.5)	0.02	

dOR = [sensitivity /(1-specificity)] / [specificity /(1-sensitivity)]

* dOR for age represent the odds ratio for each category using the category < 30 years as the reference category.

1. improvement in prediction p = 0.008
2. improvement in prediction p = 0.02
3. improvement in prediction p = 0.02 (from hysteroscopy model) and p = 0.009 (from ultrasound model)

Goodness of fit: As the objective of the exercise was to test for the feasibility of such a model and no clinical implications were recommended, the goodness of fit for the various models were not given.

Figure 7.1: Improvement in diagnostic accuracy achieved by addition of tests to the information provided by Patient history. Area under receiver operating characteristic (ROC) curves of the diagnostic models based on Patient history, patient history plus endometrial thickness measured by ultrasound scan and patient history plus ultrasound scan and hysteroscopy

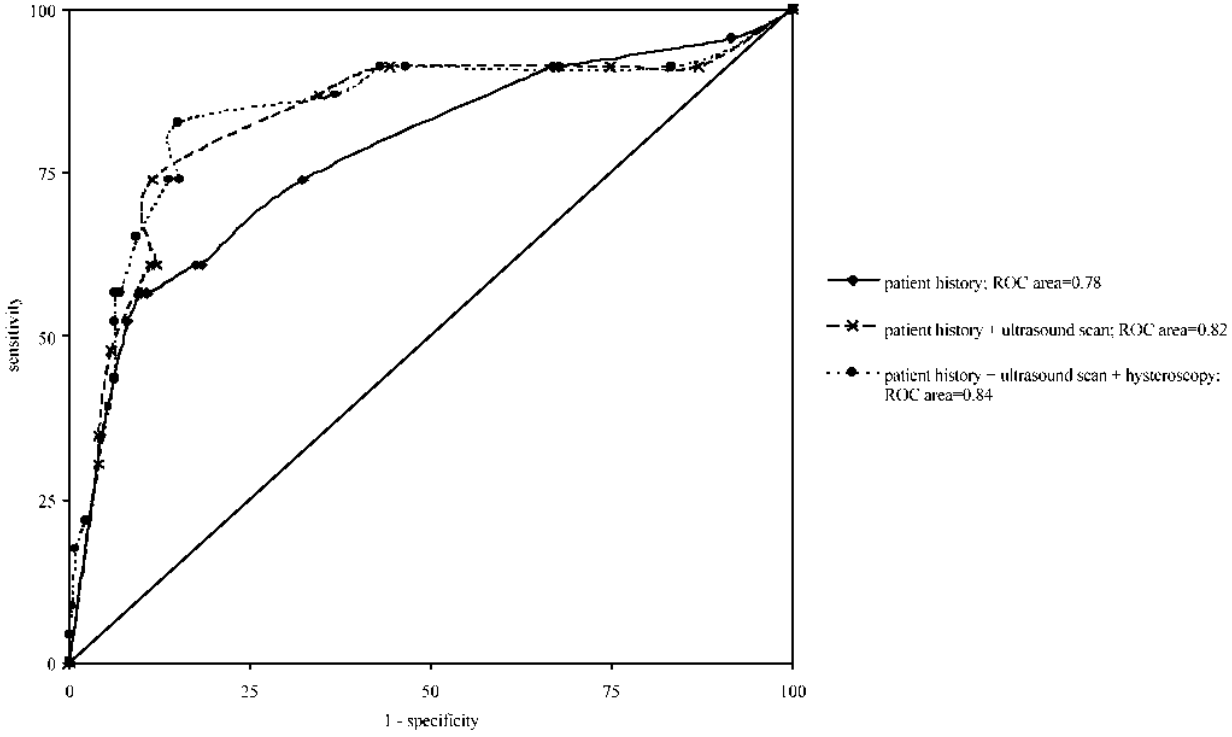
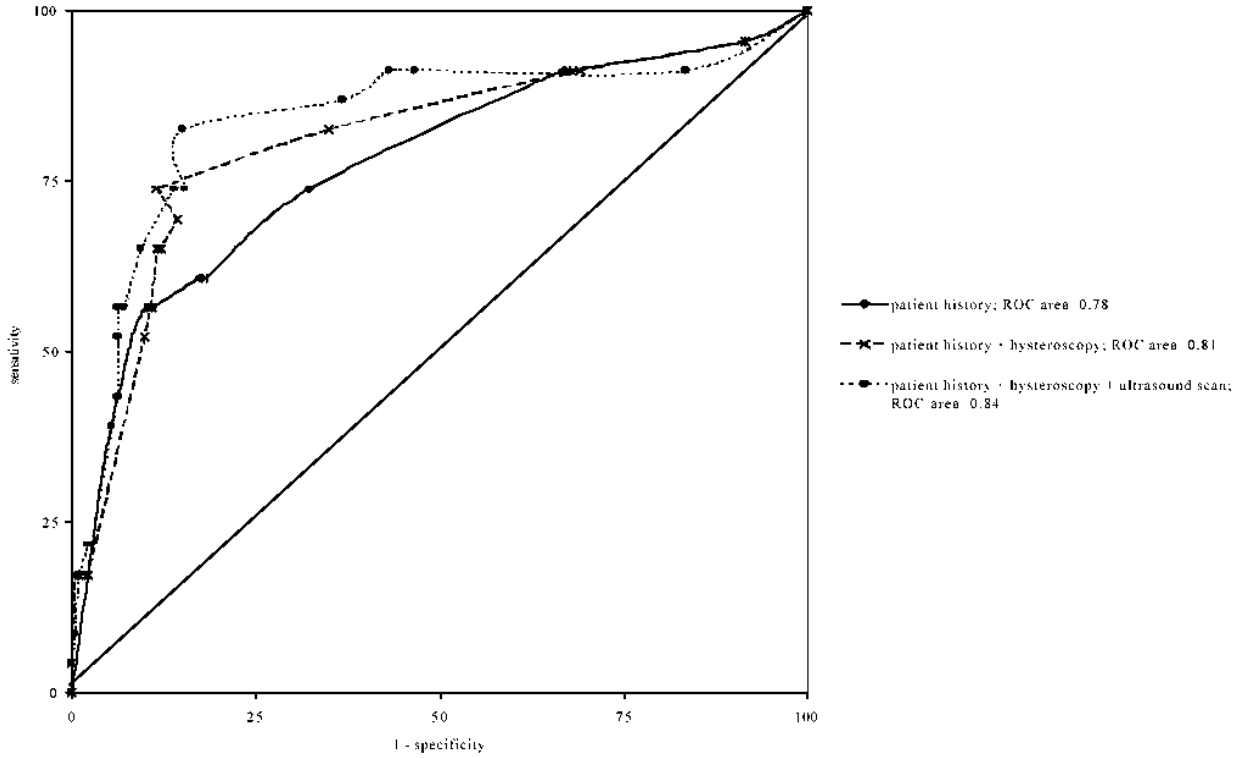


Figure 7.2: Improvement in diagnostic accuracy achieved by addition of tests to the information provided by patient history. Area under receiver operating characteristic (ROC) curves of the diagnostic models based on patient history, patient history plus hysteroscopy and patient history plus hysteroscopy and endometrial thickness measured by ultrasound scan



7.4 Discussion

7.1.1 Principal findings of the study

This study shows that multivariable analysis for evaluation of combinations of tests is feasible. In this study, univariable analyses of the various diagnostic variables had suggested potential value in history and tests for predicting serious endometrial disease (hyperplasia or cancer) in women with abnormal uterine bleeding. Multivariable analyses showed that both ultrasonography and hysteroscopy increased the prediction of serious endometrial pathology above that predicted from clinical history alone. The use of both ultrasonography and hysteroscopy together marginally (but statistically significantly) increased the predictive ability further.

7.4.2 Appraisal of quality of the study

Diagnostic tests are evaluated for accuracy in relation to a reference standard, often in isolation from the clinical context in which they will be used. Estimates of diagnostic accuracy derived in this way can lead to erroneous inferences and may artificially inflate the value of diagnostic tests (Begg 1989; Moons 1996). This study shows that to avoid misleading clinical inferences multivariable regression models can be constructed to reflect clinical practice.

This type of multivariable analysis to delineate the significance of diagnostic variables is important in facilitating meaningful clinical interpretation of tests. This is because it allows the added value of tests to be determined in light of information already available to the clinician from the history, thereby reflecting the real clinical situation. If, however, multivariable regression analysis does not reveal any additional value of a diagnostic intervention in predicting a particular disease, then it should not be used in routine clinical practice.

The clinical performance of a test can be described in terms of diagnostic accuracy, or the ability to correctly classify subjects into clinically relevant subgroups. ROC plots provide a pure index of accuracy by demonstrating the limits of a test's ability to discriminate between alternative states of health over the complete spectrum of operating conditions. Furthermore, ROC plots occupy a central or unifying position in the process of assessing and using diagnostic tools (Zweig and Campbell 1993). Quantitative ROC analysis was employed to develop a multivariable approach to determine the added value of hysteroscopy and ultrasonography tests in routine practice where some diagnostic information is already available from the clinical history.

The strength of this study and the credibility of its findings depend on the strength of the methodology. This is made clear by applying the Standard of Quality for Reporting Logistic Regression Models (Khan 1999) the objective of which is to to identify problems affecting the accuracy and interpretation of this approach to multivariable statistical analyses. Table 7.2.

Table 7.2: Application of the Standard of Quality for Reporting Logistic Regression Models

Quality Feature	Description of the Criteria	Compliance of This Study
Did the model appear to be correct?		
• Research hypothesis	It should be reported clearly	Yes
• Reporting of the variable selection process	It is adequate if it is based on consideration of the clinical importance of various factors	Yes
• Statistical significance threshold	The threshold of statistical significance for inclusion or deletion of variables (e.g. in stepwise forward or backward regression) should be reported	Yes
• Conformity to linear gradient	An attempt should be made to detect this problem, e.g. by comparison of the observed and the predicted values for the outcome over the ordinal zones or by an alternative analysis using cross-stratification	Yes
How well did the overall model work?		
• Testing for goodness-of-fit	The accuracy with which the regression model describes the data should be assessed	Yes
How important was each of the independent variable?		
• Unit of measurement	The change in dependent variable associated with a unit change in the independent variable should be reported	Yes
• Coding of variables	A description of variables if they were on an interval, ordinal or binary scale should be provided	Yes
• Testing for interactions	when the impact of one variable is dependent on the level of another variable, interactions between independent variables should be examined	Yes

The results for endometrial thickness measurement by ultrasound scan and hysteroscopy should however, be interpreted in the context of hypothesis generation. This is because the approach in this paper was limited to assessing the feasibility of developing an analytical strategy using standard software. The results from this data set of 248 women with abnormal uterine bleeding should not therefore be seen as assessing diagnostic value of the various tests. In addition, because of relative paucity of outcome events, the estimates of accuracy may be unstable, a situation in which exact methods for performing logistic regression may be more appropriate (Mehta 1995).

7.4.3 Relation to other studies considering differences in quality and results

There is no other similar study to use for discussion.

7.4.4 Practical implications for clinicians

There are no practical implications as this is a feasibility or pilot analysis.

7.4.5 Implications on research

- The findings of this chapter require confirmation in a large study.
- Diagnostic evaluation should consider multivariable analysis to evaluate combination of tests.

CHAPTER 8

DISCUSSION AND CONCLUSIONS

The proposed structure by Docherty and Smith (Docherty, Smith 1999) for the discussion of scientific papers that was used in each chapter, will be followed in this chapter as well in the following order:

- Statement of principal findings
- Appraisal of quality of the study: Strengths and weaknesses of the study
- Relation to other studies considering differences in quality and results
- Meaning of the study: possible mechanisms and implications for clinicians or policymakers
- Unanswered questions and future research

8.1 The principle findings of this thesis: (Table 8.1)

1. Hysteroscopy is accurate in ruling in endometrial cancer and hyperplasia. (See 8.3 for the clinical implications).
2. Ultrasound is accurate in ruling out endometrial cancer and hyperplasia. (See 8.3 for the clinical implications).
3. Insufficient sample on outpatient endometrial biopsy may be considered a substitute to the absence of pathology provided the hysteroscopic and sonographic endometrial assessment is consistent with endometrial atrophy. (See 8.3 for the clinical implications).
4. The rate of hyperplasia in the polyps group is approximately double that in the non-polyp group while the cancer rate was no different. (See 8.3 for the clinical implications).

5. Among the various risk factors associated with endometrial polyps, tamoxifen is associated with endometrial polyps but hormone replacement therapy is not. (See 8.3 for the clinical implications).
6. The true clinical value of a test lies in the added information over and above what was already known from the history and examination. It is feasible to develop a stepwise multivariable analytic approach to explore the added value of tests (hysteroscopy or ultrasonography) over and above history when predicting endometrial hyperplasia or cancer. This analytic strategy needs to be applied in larger datasets to draw clinical conclusions.

Table 8.1: Summary of the thesis results

Chapter 2	**Accuracy of hysteroscopy in diagnosing endometrial hyperplasia /cancer**	**Estimate of Accuracy/Association**	**Change in Probability: Pre-test →Post test**
	• Negative hysteroscopy	LR= 0.7	(Premen) 3.0%→2.1%
	•		(Postmen) 12.5%→9.1%
	• Positive hysteroscopy	LR=51.1	(Premen) 3.0%→61.2% (Postmen) 12.5%→87.9%
Chapter 3	**Accuracy of ultrasound scan in diagnosing endometrial hyperplasia /cancer**		**Change in Probability: Pre-test →Post test**
	• ET≤4mm	LR= 0.14	15%→2.4%
	• ET>9mm	LR =3.3	15%→36.8%
Chapter 4	**Factors associated with Insufficient endometrial sampling**		**Statistical Significance**
	• Ultrasonographic ET>5mm	OR= 0.19	P value=0.001
	• Hysteroscopic endometrial atrophy	OR= 4.79	P value=0.04
Chapter 5	**Pathology in endometrial polyps**		**Rates in polyps vs non polyps**
	• Hyperplasia	OR= 2.1	6/62 (9.7%) vs 9/186 (4.8%)
	• Cancer	OR= 1.5	3/36 (4.8%) vs 6/186 (3.2%)
Chapter 6	**Factors associated with endometrial polyps**		**Statistical Significance**
	• HRT	OR= 1.48	P value=0.32
	• Tamoxifen	OR= 11.21	P value=0.0009
Chapter 7	**Multivariable analysis**		**Statistical Significance**
	Clinical history + Hysteroscopy (suspicious vs normal)	ROC area= 0.81	P value=0.008 for improvement
	Clinical history + Ultrasound Scan (5mm cut-off)	ROC area= 0.82	P value=0.02 for improvement

LR: Likelihood Ratio
Premen: premenopausal
Postmen: postmenopausal
ET: Endometrial thickness
OR: Odds Ratio
ROC: Receiver- Operating Characteristic

8.2 Appraisal of quality of the study

The strength of the studies presented in this thesis and the credibility of their findings depend on the strength of the methodology. This is made clear by adhering to the checklist proposed by the STARD group (Standards for Reporting of Diagnostic Accuracy): the objective of which is to improve the quality of reporting of studies on diagnostic accuracy allowing the detection of the potential for bias in the study and to evaluate the generalisability of the results. Most of the criteria required for a high quality diagnostic test study were fulfilled by this thesis.

There were independent assessments of the test and the gold standard. The gold standard histology was performed on all patients regardless of hysteroscopic findings. There was an adequate spectrum of patients both at low and at high risk of premalignant/malignant lesions. In addition, meaningful measures of diagnostic accuracy have been used to readily enable the reader to assess how useful the test might be in practice.

Table 8.2: Application of STARD checklist to the study reported in this thesis

Section and topic	Description of the criteria	Compliance in the thesis
METHODS		
Participants:	Based on presenting symptoms, the participants had received the index test(s) and the reference standard?	Yes
Study population Setting/ Location Recruitment:		
Participant sampling:	Consecutive series of patients defined by selection criteria	Yes
Data collection:	Participants identified and data collected before the index test(s) and reference standards were performed (prospective study)	Yes
Reference standard	and its rationale	Yes
Test methods	Technical specification of Given material and methods	Yes
	Definition and rationale for the units, cut-offs and/or categories of the results of the index test(s) and the reference standard	Yes
	The number, training and expertise of the persons (a) executing and (b) reading the index test(s) and the reference standard	Yes
	Methods for calculating measures of diagnostic accuracy or making comparisons, the statistical methods used to quantify uncertainty (e.g. 95% CI) and methods for calculating test reproducibility	Yes
RESULTS		
Participants	When study was done, including beginning and ending dates of recruitment	Yes
	Clinical and demographic characteristics (e.g. age, sex, spectrum of presenting symptom(s), co morbidity, current treatment(s), recruitment centre)	Yes
Test results	A cross tabulation of the results of the index test(s) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard	Yes
	Indeterminate results, missing responses and outliers of index test(s) stratified by reference standard result and how they were handled	Yes
Estimation	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% CI)	Yes
	Estimates of variability of diagnostic accuracy between subgroups of participants (Post vs premenopausal)	Yes

The methodology was fully explained. Several accuracy indices were used, including:

- Sensitivity and specificity
- The area under the receiver operating characteristic curve
- And several categories of results were provided in the form of likelihood ratios, post-test probability of disease and Bayes' theorem (Daya 1996).

The evidence generated in this thesis, and presented in section 8.3, is graded using the scheme below for classification of evidence levels for primary research on diagnostic accuracy (Khan 2001). (Table 8.3).

Table 8.3: Classification of level of evidence for primary research on diagnostic accuracy

Level of Evidence	Description
Ia	Independent, blind comparison with reference standard among an appropriate population of consecutive patients
Ib	An independent, blind comparison with reference standard among an appropriate population of non-consecutive patients or confined to a narrow population of study patient.
IIa	An independent, non-blind comparison with reference standard among an appropriate population of consecutive patients
IIb	An independent, non-blind comparison with reference standard among an appropriate population of non-consecutive patients or confined to a narrow population of study patients.
III	An independent, blind comparison among an appropriate population of patients, but reference standard not applied to all study patients
IV	Reference standard not applied independently or expert opinion with no explicit critical appraisal, based on physiology, bench research or first principles

However the weaknesses of the thesis are:

Firstly, regarding the accuracy of the index text (Ultrasonography), the scan studies were performed by the same clinician who performed the hysteroscopy. Blindness to the result of the ultrasound scan was therefore not possible. Because of the small sample of participants who had the ultrasound scan study, additional data from another minimal access unit were incorporated to make the analysis more meaningful. This fact may be considered as one of the study's strengths as it uses multicentre data.

And secondly, one has to be careful in interpreting the findings since the study population was a selected group of women who were referred to the special clinic for investigation of abnormal uterine bleeding. This means that the generalization of our results is limited to symptomatic women. Also, as the sample was small, further work is needed on larger numbers taking into consideration the sample size required to achieve an agreed power of the test.

8.3 Practical implications for clinicians

The findings of this thesis, in my judgement, lead to the following practical recommendations:

- Firstly: hysteroscopy is highly accurate and thereby clinically useful, in diagnosing endometrial cancer in women with abnormal uterine bleeding, however the diagnosis of malignancy should continue to be positively confirmed on histology because of the major surgical treatment intervention it carries with. Although this conclusion might mean that there has been no change in clinical practice, this is a primary diagnostic test accuracy study whereby abnormal hysteroscopy was found to be accurate.

  Secondly: endometrial sampling should be performed with negative hysteroscopy, as negative hysteroscopy is not conclusive for excluding malignancy.

  (Evidence level Ia)

- Endometrial sampling may safely be skipped in women with postmenopausal bleeding, if the sonographic endometrial thickness is ≤4.0mm. This became the policy in both settings (MAST and GMAU centres) as a result of this study allowing to discharge these women without further invasive intervention (i.e. endometrial sampling or hysteroscopy).

  (Evidence level Ia)

- In postmenopausal bleeding, it is safe to disregard insufficient endometrial sample and consider it benign if sonographic endometrial thickness and hysteroscopy are normal. And yes we should be confident to reassure and discharge such a patient without the need to expose her to further investigation which conventionally means inpatient dilatation and curettage. This policy became the standard practice in the One-stop clinic as a result of this study.

  (Evidence level Ia)

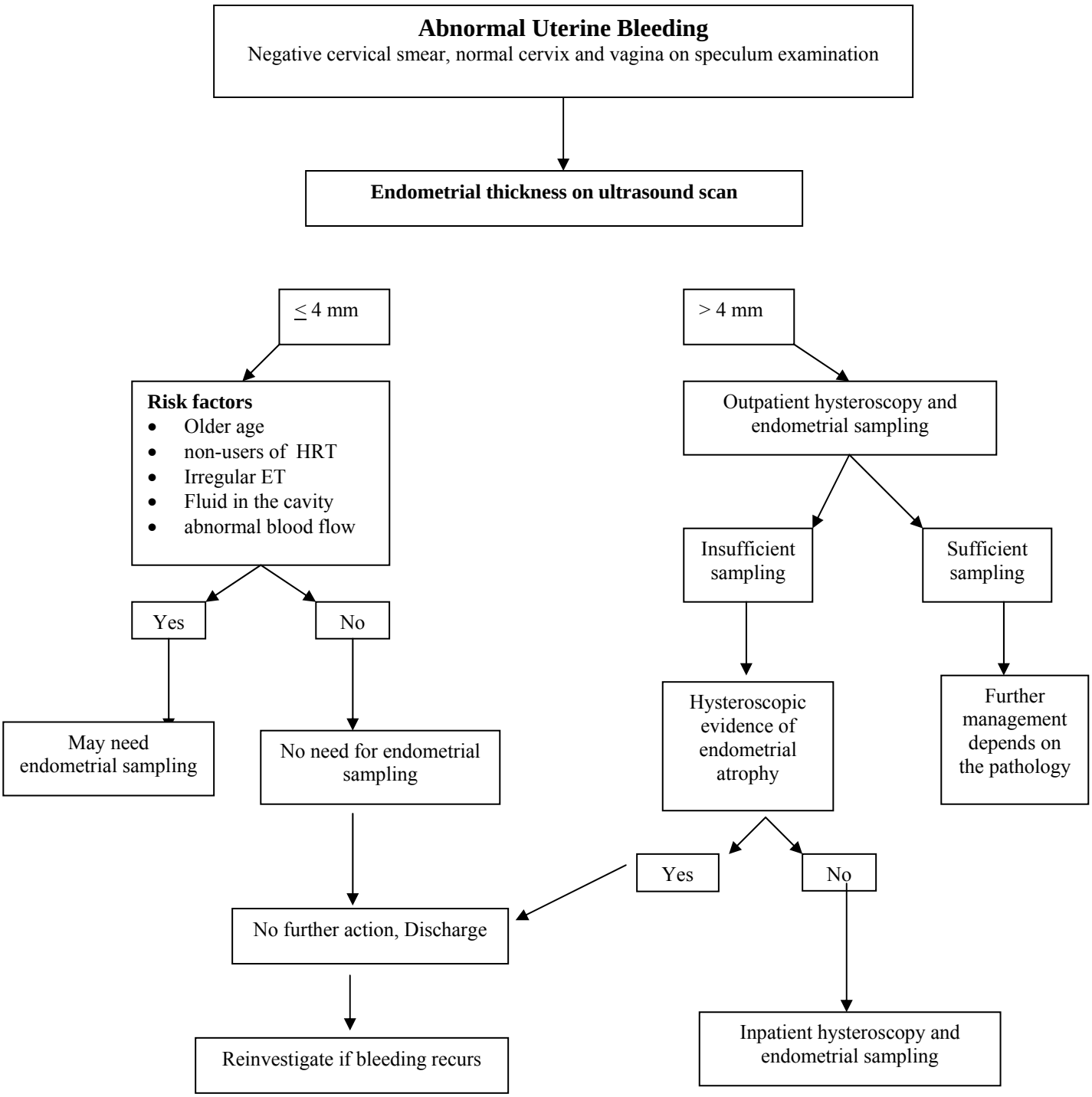
- As this study concluded that the rate of hyperplasia in the polyps group is approximately double that in the non-polyp group, the clinical implication recommends that the detection of polyps should lead to their removal followed by histological assessment as there is some evidence indicates that the risk of recurrence of endometrial polyp is expected in the presence of an abnormal proliferative activity or hyperplasia both in the endometrial polyp and the surrounding endometrium (Bouda J Jr 2000). However one can be reassured that the risk of cancer is not increased in the polyps group.

  (Evidence level Ia)

- There are numerous recognized reasons for long-term non-compliant use of HRT; but the risk of HRT inducing endometrial polyps should not be one of these. It should be always remembered that tamoxifen treatment is a significant risk factor associated with endometrial polyps. Therefore women on tamoxifen therapy who present with abnormal uterine bleeding need prompt investigation. (Evidence level Ia)

Based on the principles discussed above the following is a flow chart on how to manage women with abnormal uterine bleeding:

Flow Chart: Management of women with abnormal uterine bleeding



Implications on research

This thesis, in my view, has the following recommendations for future research:

- To determine the contribution of various tests used in women with abnormal uterine bleeding, further studies using larger data sets should be carried out using the analytical technique (stepwise analysis) outlined in chapter 7.
- The effectiveness and cost- effectiveness of the available diagnostic tools (hysteroscopy /ultrasound /biopsy) are required using modeling.
- The management of polyps requires randomised trials to address the effectiveness of various methods of treatment.
- Based on the methodology shown in this feasibility study, systematic reviews using Individual Patients Data of relevant or large primary studies are needed to evaluate the accuracy of combinations of tests used in the ambulatory assessment of menstrual disorders
- Guideline or national consensus on how to investigate menstrual disorders/abnormal uterine bleeding is needed. Such a guideline was recently published in Scotland (the Scottish Intercollegiate Guideline Network 2002). The guidelines are helpful, not only to guide practice, but to establish the current state of knowledge and can be updated as more research information becomes available.

APPENDIX 1

Letters from the Chief of Research and Development Committee and from the Clinical Governance Manager regarding the ethical approval

THE UNIVERSITY
OF BIRMINGHAM

Division of Reproductive
and Child Health
Academic Department of
Obstetrics and Gynaecology

Birmingham Women's Hospital
Edgbaston
Birmingham B15 2TG
United Kingdom
Telephone 0121 627 2775
Fax 0121 415 4857

Head of Division and Department
Professor Martin Whittle
Mobile 0[illegible]
Email M.J.Whittle@[illegible]

Divisional Manager
Mrs L. Baillie
Email [illegible]

MJW/MW

30th August, 2002.

TO WHOM IT MAY CONCERN

MD Evaluation of Ambulatory Diagnosis of Menstrual Disorders
Dr. Shagaf Bakour

This study was commenced in 1996, at which time no particular ethical issues were considered to exist. The study itself looked at three methods of evaluation and these were used in each of the patients included in the study whenever possible. In fact the three methods used were all standard of care and the patient had the benefit of all three procedures. A comparison was then made as to the value of each individual method.

Since that time there has been a further development in the Declaration of Helsinki but this became active in October 2000 and, therefore, post-dated the onset of this particular study.

We consider that the study, when it was established, did not require ethical approval.

Yours faithfully,

M.J. Whittle,
Professor.

Fetal Medicine: Professor M. J. Whittle, 0121 627 2775
Ireland: Mr M. Kilo, 0121 627 2779
Gynaecological Oncology: Professor D. Lindley, 0121 623 6833
Senior Lecturer: Mr S Kelow, 0121 623 6834
Reproductive Physiology: Non-Clinical Professor, Professor A. Ahmed, 0121 627 2709
Obstetrics and Gynaecology: Senior Lecturer, Mr J. Gupta, 0121 [illegible]

Birmingham Women's Health Care NHS

NHS Trust

POSITIVE ABOUT PEOPLE

Shagaf Bakour
Specialist Registrar
Academic Department of
Obstetrics and Gynaecology
Birmingham Women's Hospital
Birmingham
B15 2TG
UK

Malcolm Bowcock
Clinical Governance & Audit Manager
Birmingham Women's Hospital
Metchley Park Road
Edgbaston
Birmingham
B15 2TG
UK

Switchboard: 0121 472 1377
DDI 0121 623 6944

25th March 2002

Dear Shagaf

I, as Clinical Governance & Audit Manager of Birmingham Women's Health Care NHS Trust, confirm that approval for the thesis "Evaluation of Ambulatory Diagnosis of Menstrual Disorders" by Dr Bakour, was not required by this institution.

Yours sincerely

MBowcock

Malcolm Bowcock
Clinical Governance & Audit Manager

Chairman – Anne Owen Chief Executive – Caroline Wigley

APPENDIX 2

Clarification about author's contribution to chapter 3

The clinical work presented in this thesis through all chapters, was carried out at Birmingham women's hospital by the effort of the author; this included the 248 patients/data from the One-stop clinic.

Chapter 3, however, presented data from the above cohort as well as data from another centre: the Gynaecology Minimal Access Unit, Ninewells Hospital, Dundee. This included 52 patients/data. This chapter appeared as a peer review article whereby combined data from both centres were presented. To keep the consistency, it is been decided to keep Dundee's data when writing this thesis.

REFERENCES

Standards for Reporting of Diagnostic Accuracy (STARD group). (http://www.consort-statement.org/append).

52nd WMA General Assembly. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2000; 284, 3043-3049.

Achiron R, Lipitz S, Sivan E, Goldenberg M, Mashiach S. Sonohysterography for ultrasonographic evaluation of tamoxifen- associated cystic thickened endometrium. *J Ultrasound Med* 1995; 14:685-88.

Al-Azzawi F. Hysteroscopy or ultrasound? Curr Opin Obstet Gynecol 8, 246-249. 1996.

Aleem F, Predanic M, Calame R, Moukhtar M, Pennisi J. Transvaginal color and pulsed Doppler sonography of the endometrium: a possible role in reducing the number of dilatation and curettage procedures. *J Ultrasound Med* 1995; 14:139-45.

Antoni J, Folch E, Costa J et al. Comparison of cytospat and pipelle endometrial biopsy instruments. *Eur J Obstet Gynecol Reprod Biol* 1997; 72:57-61.

Bakour S H, Khan K S, Gupta J K, Newton J R. The diagnostic accuracy of ultrasound scan in predicting endometrial hyperplasia or cancer in postmenopausal bleeding. *Acta Obstet Gynecol Scand* 1999; 78(5): 447-451.

Bakour SH, Dwarakanath LS, Khan KS, Newton JR. The diagnostic accuracy of outpatient miniature hysteroscopy in predicting premalignant and malignant endometrial lesions. *Gynaecol Endosc* 1999; 8(3): 143-148.

Bakour SH, Khan KS, Gupta JK. The risk of premalignant and malignant pathology in endometrial polyps. *Acta Obstet Gynecol Scand* 2000; 79:317-20.

Begg CB. Experimental design of medical imaging trials: issues and opinions. *Invest Radiol* 1989; 24: 934-926.

Ben Baruch G, Seidman DS, Schiff E, Moran O, Menczer J. Outpatient endometrial sampling with the Pipelle curette. *Gynecol Obstet Invest* 1994; 37:260-262.

Ben Yehuda OM, Kim YB, Leuchter RS. Does hysteroscopy improve upon the sensitivity of dilatation and curettage in the diagnosis of endometrial hyperplasia or carcinoma? *Gynecol Oncol* 1998; 68:4-7.

Bender R, Grouven U. Ordinal logistic regression in medical research. *J R Coll Physicians Lond* 1997; 31:546-51.

Bouda J Jr, Hradecky L, Rokyta Z. Hysteroscopic polypectomy versus fractionated curettage in the treatment of corporal polyps--recurrence of corporal polyps. *Ceska Gynekologie* 2000; 65(3):147-51. EBM Reviews - Cochrane Central Register of Controlled Trials.

Brennan P, Croft P I. nterpreting the results of observational research: chance is not such a fine thing. *Br Med J* 1994; 309: 727-730.

Cheng W-F, Lin H-H, Trong P-L, Huang S-C. Comparison of endometrial changes among symptomatic tamoxifen-treated and nontreated premenopausal and postmenopausal breast cancer patients. *Gynaecol Oncol* 1997; 66: 233-237.

Chien PFW, Khan KS, Ogston S, Owen P. The Diagnostic Accuracy of Cervico-vaginal foetal Fibronectin in predicting Preterm delivery - An Overview. *Br J Obstet Gynaecol* 1997; 104: 436-444.

Chien PF, Khan KS. Evaluation of a clinical test. II: Assessment of validity. *British Journal of Obstetrics & Gynaecology* 2001; 108:568-72.

Choo YC, Mak KC, Hsu C, Wong TS, Ma HK. Post-menopausal uterine bleeding of nonorganic cause. *Obstet Gynecol* 1985; 225-231.

Cicinelli E, Comi N, Scorcia P, Petrizzo O, Epifani S. Hysteroscopy for diagnosis and treatment of endometrial adenocarcinoma precursors. *Eur J Gynaec Oncol* 1993; 5: 425-436.

Clark T J, Mann C H, Shah N, Khan K S, Song F, Gupta J K. Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: a systematic quantitative review. *Br J Obstet Gynaecol* 2002a; 109: 313-321.

Clark TJ, Voit D, Gupta JK, Hyde C, Song F, Khan KS. Accuracy of hysteroscopy in the diagnosis of endometrial cancer and hyperplasia: a systematic quantitative review. *JAMA* 2002b; 288(13):1610-1621.

Clark TJ, Mann CH, Shah N, Khan KS, Song F, Gupta JK. Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial hyperplasia. *Acta Obstet Gynecol Scand* 2001; 80:784-93.

Cohen I, Rosen DJD, Shapira J, Cordoba M, Gilboa S, Altaras MM, Yigael D, Beyth Y. Endometrial changes in postmenopausal women treated with tamoxifen for breast cancer. *Br J Obstet Gynaecol* 1993; 100: 567-570.

Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med* 1993; 118: 201-210.

Coope J, Marsh J. Can we improve compliance with long-term HRT? *Maturitas* 1992; 15:151-58.

Coulter A, Kelland A, Long Aea. The management of menorrhagia. *Effective Health Care Bulletin* 1995; 9. 9. 1995.

de Jong P, Doel F, Falconer A. Outpatient diagnostic hysteroscopy. *Br J Obstet Gynaecol* 1990; 97:299-303.

Deeks JJ, Morris JM. Evaluating diagnostic tests. *Baillière=s Clinical Obstetrics & Gynaecology* 1996; 10: 613-630.

Department of clinical epidemiology and biostatistics. How to read clinical journals. II. To learn about a diagnostic test. *Can Med Assoc J* 1981; 124: 703-710.

Dijkhuizen FP, De Vries LD, Mol BW et al. Comparison of transvaginal ultrasonography and saline infusion sonography for the detection of intracavitary abnormalities in premenopausal women. *Ultrasound Obstet Gynecol* 2000; 15:372-76.

Dijkhuizen FP, Mol BW, Brolmann HA, Heintz AP. The accuracy of endometrial sampling in the diagnosis of patients with endometrial carcinoma and hyperplasia: a meta-analysis. *Cancer* 2000; 89:1765-72.

Docherty M, Smith R. The case for structuring the discussion of scientific papers. [see comments]. *BMJ* 1999; 318:1224-25.

Dorum A, Kristensen G B, Langebrekke A, Sornes T, Skaar O. Evaluation of endometrial thickness measured by endovaginal ultrasound in women with postmenopausal bleeding. *Acta Obstet Gynecol Scand* 1993; 72: 116-119.

Dounes E, Al-Azzawi F. How well do perimenopausal patients accept outpatient hysteroscopy? Visual analogue of acceptability and pain in 100 women. *Eur J Obstet Gynecol Reprod Biol* 1993; 48, 37.

Dubimsky T J, Parvey Hr, Maklad N. The role of transvaginal sonography and endometrial biopsy in the evaluation of peri- and postmenopausal bleeding. *A J R* 1997; 169: 145-149.

Exacoustos C, Zupi E, Chiaretti M, Arduini D, Romanini C. Endometrial evaluation in postmenopausal breast cancer patients receiving tamoxifen: an ultrasound, color flow Doppler, hysteroscopic and histological study. *Ultrasound Obstet Gynecol* 1995; 6: 435-442.

Farquhar CM, Lethaby A, Sowter M, Verry J, Baranyai J. An evaluation of risk factors for endometrial hyperplasia in premenopausal women with abnormal menstrual bleeding. *Am J Obstet Gynecol* 1999;181:525-29.

Fineberg HV, Hiatt HH. Evaluation of medical practices: The case for technology assessment. *N Engl J Med* 1979; 301: 1086-1091.

Finikiotis G. Hysteroscopy: A review. *Obstet Gynecol Surv* 1994; 49: 273-283.

Fleisher AC, Kalemeris GC, Machin JE, Entman SS, James E. Sonographic depiction of normal and abnormal endometrium with histopathologic correlation. *J Ultrasound Med* 1996; 5: 445-452.

Fraser IS, Baird DT. Endometrial cystic glandular hyperplasia in adolescent girls. *Journal of Obstetrics & Gynaecology of the British Commonwealth* 1972; 79:1009-15.

Gallop DG, Stock RJ. Adenocarcinoma of the Endometrium in Women 40 Years of Age or Younger. *Obstet Gynecol* 1984; 64: 417-420.

Gimpleson RJ. Panoramic hysteroscopy with directed biopsies vs dilatation and curettage for accurate diagnosis. *J Reprod Med* 1984; 29: 8-12.

Gimpleson RJ, Rappold HO. A comparative study between panoramic hysteroscopy with directed biopsies and dilatation and curettage. *Am J Obstet Gynaecol* 1998; 158: 489-492.

Giusa- Chiferi M G, Goncalves WJ, Baracat E C, Cavalcanti de Albuquerque, Neto L, Bortrletto C C R, Rodrigues de Lima G. Transvaginal ultrasonound, uterine biopsy and hysteroscopy for postmenopausal bleeding. *Int J Gynecol Obstet* 1996; 55: 39-44.

Goodman A. Abnormal genital tract bleeding. *Clin Cornerstone* 2000; 3:25-35.

Gordon SJ, Westgate J. The incidence and management of failed Pipelle sampling in a general outpatient clinic. *Aust N Z J Obstet Gynaecol* 1999;39:115-18.

Granberg S, Wikland M, Karlsson B, Friberg L. Endometrial thickness by endovaginal ultrasonography for identifying endometrial abnormality. *Am J Obstet Gynecol* 1991; 164: 47-52.

Granberg S, Ylostalo P, Wikland M, Karlsson B. Endometrial sonographic and histologic findings in women with and without hormonal replacement therapy suffering from postmenopausal bleeding. *Maturitas* 1997; 27: 35-40.

Greenhalgh T. How to read a paper. Papers that report diagnostic or screening tests. *BMJ* 1997; 315:540-543.

Guido RS, Kanbour-Shakir A, Rulin MC, Christopherson WA. Pipelle endometrial sampling: sensitivity in the detection of endometrial cancer. *J Reprod Med* 1995; 40: 553-555.

Gupta JK, Chien PF, Voit D, Clark TJ, Khan KS. Ultrasonographic endometrial thickness for diagnosing endometrial pathology in women with postmenopausal bleeding: a meta-analysis. *Acta Obstet Gynecol Scand 2002*; 81(9):799-816.

Gupta JK, Wilson S, Desai P, Hau C. How should we investigate women with postmenopausal bleeding? *Acta Obstet Gynecol Scand* 1996; 75:475-79.

Haller H, Matejcic N, Rukavina B, Krasevic M, Rupcic S, Mozetic D. Transveginal sonography and hysteroscopy in women with postmenopausal bleeding. *Int J Gynecol Obstet* 1996; 54: 155-159.

Hanggi W, Bersinger N, Altermatt HJ, Birkhauser MH. Comparison of transvaginal ultrasonography and endometrial biopsy in endometrial surveillance in postmenopausal HRT users. *Maturitas* 1997; 27(2):133-43.

Hanley JA, McNeil BJ. The meaning and use of the area under a receiver-operating (ROC) curve. *Radiology* 1982; 143: 29-36.

Hardell L. Tamoxifen as risk factor for carcinoma of corpus uteri (letter). *Lancet* 1988; 3: 563.

Hunter DC, McClure N. Abnormal uterine bleeding: an evaluation endometrial biopsy, vaginal ultrasound and outpatient hysteroscopy. *Ulster Med J* 2001; 70:25-30.

Irwig L, Macaskill P Glasziou P Fahey M. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol 1995; 48: 119-130.

Ismail MS. The effects of tamoxifen on the uterus. *Current Opinion in Obstetrics and Gynaecology* 1996; 8: 27-31.

Jaeschke R, Guyatt G, Sackett DL. For the Evidence-based Medicine working group. Users guide to the medical literature III. How to use an article about a diagnostic test (2 parts). *JAMA* 1994; 271:389-391, 703-707.

Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994; 271:389-91.

Jeffery JD, Taylor R, Robertson DI, Stuart GC. Endometrial carcinoma occurring in patients under the age of 45 years. *Am J Obstet Gynecol* 1987; 156:366-70.

Karlsson B, Granberg S, Wikland M, Ylostalo P, Torvid K, Marsal K. et al. Transvaginal ultrasonography of the endometrium in women with postmenopausal bleeding- A Noradic multicenter study. *Am J Obstet Gynecol* 1995; 172: 1488-1494.

Khan K, ter Riet G, Popay J, Nixon J Kleijnen J. Study quality assessment. In: Khan KS, Ter Riet G, Glanville J, Sowden AJ, Kleijnen J (Eds). Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for Carrying Out or Commissioning Reviews. 2nd Edition. (ISBN 1900640201). CRD Report No.4.2001.York: NHS Centre for Reviews and Dissemination (CRD), University of York, 2001.(URL: http://www.york.ac.uk/inst/crd/report4.htm). 2001.

Khan KS, Chien PFW, Dwarakanath LS. Logistic regression models in Obstetrics and Gynaecology literature. *Obstet Gynecol* 1999 93(6), 1014-1020.

Khan KS, Dinnes J, Kleijnen J. Systematic reviews to evaluate diagnostic tests. *Eur JObstet Gynecol Reprod Biol* 2001; 95(1): 6-11.

Khan KS, Chien PF. Evaluation of a clinical test. I: assessment of reliability. *Br J Obstet Gynaecol* 2001; 108:562-67.

Kufahl J, Pedersen I, Sindberg Eriksen P, Helkjaer LG, Jensen KL, de Nully O et al. Transvaginal ultrasound, endometrial cytology sampled by Gynoscann and histology obtained by Uterine Explora Curette compared to the histology of the uterine specimen. A prospective study in pre- and postmenopausal women undergoing elective hysterectomy. *Acta Obstet Gynecol Scand* 1997; 76(8): 790-796.

Kurman RJ, Kaminski PF, Norris HJ. The behavior of endometrial hyperplasia. A long-term study of "untreated" hyperplasia in 170 patients. *Cancer* 1985; 56:403-12.

Lahti E, Blanco G, Kauppila A, Apaja-Sarkkinen M, Taskinen P J, Laatikainen T. Endometria; changes in postmenopausal breast cancer patients receiving tamoxifen. *Obstet Gynecol* 1993; 81: 660-664.

Lethaby A, Farquhar C, Sarkis A, Roberts H, Jepson R, Barlow D. Hormone replacement therapy in postmenopausal women. *Cochrane Database of Systematic Reviews*. 1, 2003.

Lidor A, Ibea. Histopathologic Finding in 226 Women With Post-Menopausal Uterine Bleeding. *Acta Obstet Gynecol Scand* 1986; 65: 41-43.

Linnet K. A review on the methodology for assessing diagnostic tests. *Clin Chem* 1988; 34: 1379-1386.

Loffer FD. Hysteroscopy with selective endometrial sampling compared with D&C for abnormal uterine bleeding: the value of a negative hysteroscopic view. *Obstet Gynecol* 1989;73:16-20.

Long CA. Evaluation of patients with abnormal uterine bleeding. *Am J Obstet Gynecol* 1996; 175:784-86.

Maja H JR, Barbosa IC Farias JP Ladipo OA & Coutnho EM. Evaluation of the Endometrial Cavity During Menopause. *Int J Gynaecol Obstet* 1996; 52: 61-66.

Maja Jr H, Barbosa IC, Marques D, Calmon LC, Ladipo LC, Coutinho EM. Hysteroscopy and transvaginal sonography in menopausal women receiving hormone replacement therapy. *J Am Assoc Gynecolo Laparosc* 1996; 4: 13-18.

Marconi D, Exacoustos C, Cangi B et al. Transvaginal sonographic and hysteroscopic findings in postmenopausal women receiving tamoxifen. *J Am Assoc Gynecol Laparosc* 1997; 4:331-39.

McAlister FA, Straus SE Sackett DL. Why we need large, simple studies of the clinical examination: the problem and a proposed solution. Lancet 1999; 354: 1721-1724..

Mehta CR, Patel NR. Exact Logistic Regression: Theory and Examples. *Stat Med* 1995; 14: 2143-2160.

Mencaglia L. Hysteroscopy and adenocarcinoma. *Obstet Gynecol Clin North Am* 1995; 22:573-79.

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. [see comments]. *Lancet* 1999; 354:1896-900.

Moons CKGM. Hazards of a univariable approach in diagnostic test evaluation in Diagnostic Research Theory and Application. (PhD thesis), 37-52. 1996.

Nasri M N, Coast G J. Correlation of ultrasound findings and endometrial histopathology in postmenopausal women. *Br J Obstet Gynaecol* 1989; 96: 1333-1338.

Nasri M N, Shepherd J H, Setchell M E, Lowe D G, Chard T. The role of vaginal scan in measurement of thickness in postmenopausal women. *Br J Obstet Gynaecol* 1991; 98: 470-475.

Norman GR, Streiner DL. Screwups, oddballs, and other vagaries of science.
Locating outliers, Handling missing data, and transformations. In: Norman GR, Streiner DL, editors. Biostatistics: The bare essentials. St. Louis: Mosby; 1994. P. 202-210.

O'Connell LP, Fries MH, Zeringue E, Brehm W. Triage of abnormal postmenopausal bleeding: a comparison of endometrial biopsy and transvaginal sonohysterography versus fractional curettage with hysteroscopy. *Am J Obstet Gynecol* 1998; 178:956-61.

Owen P, Khan KS. Fetal growth velocity in the prediction of intrauterine growth retardation in a low risk population. *Br J Obstet Gynaecol* 1998; 105: 536-540.

Pal L, Lapensee L, Toth TL, Isaacson KB. Comparison of office hysteroscopy, transvaginal ultrasonography and endometrial biopsy in evaluation of abnormal uterine bleeding. *JSLS* 1997; 1:125-30.

Pickar J H, Thorneycroft I, Whitehead M. Effects of hormone replacement therapy on the endometrium and lipid parameters: a review of randomized clinical trials, 1985 to 1995. *Am J Obstet Gynecol* 1998; 178(5): 1087-1099.

Powles TJ, Bourne T Athanasiou S Chang J Grubock K Ashley S Oakes L Tidy A Davey J Viggers J Humphries S Collins W. The effect of norethisterone on endometrial abnormalities id identified by transvaginal ultrasound screening of healthy post-menopausal women on tamoxifen or placebo. *Br J Cancer* 1998; 78: 272-275.

Press Center of the World Medical Association. The World Medical Association, Inc. - Press Centre - 48th WMA General Assembly: Agenda (September 1996). www.wma.net/e/press/96_5.html . 1996.

Reid MC, Lachs MS Feinstein AR. Use of methodological standards in diagnostic tests research. Getting better but still not good. *JAMA* 1995, 274: 645-651.

Reslova T, Tosner J, Resl M, Kugler R, Vavrova I. Endometrial polyps. A clinical study of 245 cases. *Arch Gynecol Obstet* 1999; 262:133-39.

Rodriguez GC, Yaqub N King ME. A comparison of the Pipelle device and the Vabra aspirator as measured by endometrial denudation in hysterectomy specimens: the Pipelle device samples significantly less of the endometrial surface than the Vabra aspirator. *Am J Obstet Gynecol* 1993; 168: 55-59.

Rose PG. Endometrial carcinoma. *N Engl J Med* 1996;335:640-649.

Royal College of Obstetrics and Gynaecologists. The Initial Management of Menorrhagia. Evidence -Bsed Clinical Guidelines No.1 London: RCOG Press; 1998. 1.

Saidi MH, Sadler RK, Theis VD, Akright BD, Farhat SA, Villanueva GR. Comparison of sonography, sonohysterography, and hysteroscopy for evaluation of abnormal uterine bleeding. *J Ultrasound Med* 1997; 16: 587-591.

Scottish Intercollegiate Guidelines Network. Investigation of Post-Menopausal Bleeding Section 5: Interpretation of transvaginal ultrasound (TVUS). Website address: http://www.sign.ac.uk Last modified 7/10/02.

Siegler AM. Office hysteroscopy. *Obstet Gynecol Clin North Am* 1995; 22:457-71.

Smith-Bindman R, Kerlikowske K, Feldstein VA et al. Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. [See comments]. *JAMA* 1998; 280:1510-1517.

Spencer CP, Whitehead MI. Endometrial assessment re-visited. *Br J Obstet Gynaecol* 1999; 106: 623-632.

Sterne JAC, Smith GD Cox DR. Sifting the evidence what's wrong with significance tests? *BMJ* 2001; 322: 226-231.

Stock RJ, Kanbour A. Prehysterectomy Curettage. *Obste & Gynecol* 1975; 45: 537-54.

Stovall TG, Photopulos GJ Poston WM Ling FW Sandles LG. Pipelle endometrial sampling in patients with known endometrial carcinoma. *Obstet Gynecol* 1991; 77: 954-956.

Stovall TG, Ling FW, Morgan PL. A prospective, randomized comparison of the Pipelle endometrial sampling device with the Novak curette. *Am J Obstet Gynecol* 1991; 165:1287-90.
Swets J. Measuring the accuracy of diagnostic systems. *Science* 1988; 240: 1285-1293.

Terakawa N, Kigawa J, Taketani Y et al. The behavior of endometrial hyperplasia: a prospective study. Endometrial Hyperplasia Study Group. *J Obstet Gynaecol Res* 1997; 23:223-30.

Tosteson ANA, Weinstein MC, Wittenberg J, Begg CB. ROC curve regression analysis: The use of ordinal regression models for diagnostic test assessment. *Environ Health Perspec* 1994; 102: 73-78.

Towbin NA, Gviasda IM & March CM. Office Hysteroscopy Versus Transvaginal Ultrasonography in the Evaluation of Patients With Excessive Uterine Bleeding. *Am J Obstet & Gynecol* 1996; 174: 1678-1682.

Uno LH Sugimoto O, Carvalho FM Bagnoli VR Fonseca AM Pinotti JA. Morphologic hysteroscopic criteria suggestive of endometrial hyperplasia. *Int J Gynecol Obstet* 1995; 49: 35-40.

Valle RF. Hysteroscopic evaluation of patients with abnormal uterine bleeding. *Surg Gynecol Obstet* 1981; 153:521-26.

Walter SD, Jadad AR. Meta-analysis of screening data: A survey of the literature. *Statist Med* 1999; 18: 3409-3424.

Wolman I, Amster R, Hartoov J, Gull I, Kupfermintz M, Lessing JB, et al. Reproducibility of transvaginal ultrasonographic measurements of endometrial thickness in patients with postmenopausal bleeding. *Gynecol Obstet Invest* 1998; 46: 191-4.

Word B, Gravlee LC, Windeman GL. The Fallacy of Simple Uterine Curettage. *Obstet & Gynecol* 1958; 12: 642-648.

PUBLICATIONS FROM THE THESIS

1. BAKOUR SH, DWARAKANATH LS, KHAN KS, NEWTON JR. The diagnostic accuracy of outpatient miniature hysteroscopy in predicting premalignant and malignant endometrial lesions. *Gynaecol Endosc* 1999; 8 (3), 143-8.

2. BAKOUR SH, KHAN KS, GUPTA JK, NEWTON JR. The diagnostic accuracy of ultrasound scan in predicting endometrial hyperplasia or cancer in postmenopausal bleeding. *Acta Obstet& Gynecol Scand* 1999; 78(5), 447-51.

3. BAKOUR SH, KHAN KS, GUPTA JK, Controlled analysis of factors associated with insufficient sample on outpatient endometrial biopsy. *Br J Obstet & Gynaecol* 2000; 10: 1312-4.

4. BAKOUR SH, KHAN KS, GUPTA J K. Letter to Editor: Transvaginal ultrasonography and endometrial histology in peri- and postmenopausal women on hormone replacement therapy. Letter to Editor. *Br J Obstet & Gynaecol* 2000; 107(2): 295.

5. BAKOUR SH, KHAN KS, GUPTA JK. The risk of premalignant and malignant pathology in endometrial polyps. *Acta Obstet& Gynecol Scand* 2000 (Apr); 79(4): 317-20.

6. BAKOUR SH, KHAN KS, GUPTA JK. Risk factors associated with endometrial polyps in abnormal uterine bleeding. *Int J Gynecol & Obstet* 2002 Feb;76(2):165-8.

7. CLARK TJ, BAKOUR SH, KHAN KS, GUPTA JK. Evaluation of outpatient hysteroscopy and ultrasonography in the diagnosis of endometrial disease. *Obstet Gynecol*. 2002 Jun;99(6):1001-7