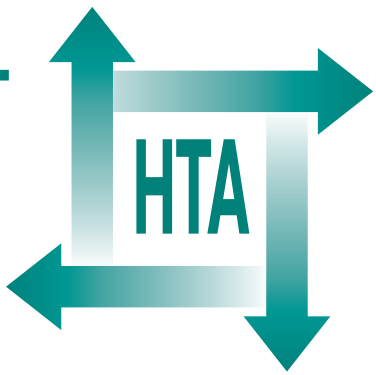
Health Technology Assessment 2004; Vol. 8: No. 34

Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status

HOD Critchley, P Warner, AJ Lee, S Brechin, J Guise and B Graham

September 2004

Health Technology Assessment
NHS R&D HTA Programme



How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status

HOD Critchley,[1*] P Warner,[2] AJ Lee,[2] S Brechin,[1] J Guise[2] and B Graham[3]

[1] Obstetrics and Gynaecology, University of Edinburgh Medical School, UK
[2] Public Health Sciences, University of Edinburgh Medical Buildings, UK
[3] Graham Consulting, Edinburgh, UK

* Corresponding author

Declared competing interests of authors: none

Published September 2004

This report should be referenced as follows:

Critchley HOD, Warner P, Lee AJ, Brechin S, Guise J, Graham B. Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status. *Health Technol Assess* 2004;**8**(34).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE.

NHS R&D HTA Programme

The research findings from the NHS R&D Health Technology Assessment (HTA) Programme directly influence key decision-making bodies such as the National Institute for Clinical Excellence (NICE) and the National Screening Committee (NSC) who rely on HTA outputs to help raise standards of care. HTA findings also help to improve the quality of the service in the NHS indirectly in that they form a key component of the 'National Knowledge Service' that is being developed to improve the evidence of clinical practice throughout the NHS.

The HTA Programme was set up in 1993. Its role is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined to include all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care, rather than settings of care.

The HTA programme commissions research only on topics where it has identified key gaps in the evidence needed by the NHS. Suggestions for topics are actively sought from people working in the NHS, the public, consumer groups and professional bodies such as Royal Colleges and NHS Trusts.

Research suggestions are carefully considered by panels of independent experts (including consumers) whose advice results in a ranked list of recommended research priorities. The HTA Programme then commissions the research team best suited to undertake the work, in the manner most appropriate to find the relevant answers. Some projects may take only months, others need several years to answer the research questions adequately. They may involve synthesising existing evidence or designing a trial to produce new evidence where none currently exists.

Additionally, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme is able to commission bespoke reports, principally for NICE, but also for other policy customers, such as a National Clinical Director. TARs bring together evidence on key aspects of the use of specific technologies and usually have to be completed within a limited time period.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this monograph was commissioned by the HTA Programme as project number 95/17/06. As funder, by devising a commissioning brief, the HTA Programme specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health.

ISSN 1366-5278

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.
Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.

Abstract

Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status

HOD Critchley,[1*] P Warner,[2] AJ Lee,[2] S Brechin,[1] J Guise[2] and B Graham[3]

[1] Obstetrics and Gynaecology, University of Edinburgh Medical School, UK
[2] Public Health Sciences, University of Edinburgh Medical Buildings, UK
[3] Graham Consulting, Edinburgh, UK
* Corresponding author

Objectives: To compare three outpatient methods of endometrial evaluation in terms of performance, patient acceptability and cost-effectiveness.
Design: Pragmatic unblinded trial randomised separately within three groups determined by risk of endometrial cancer.
Setting: The gynaecology outpatient clinic of a large city hospital in Edinburgh, Scotland.
Participants: Women referred for investigation and management of abnormal bleeding between January 1999 and May 2001.
Interventions: Investigations were: blind biopsy alone, hysteroscopy with biopsy, ultrasound evaluation including transvaginal ultrasound, and, in the low-risk group, the option of no investigation. Within this design, two devices for obtaining endometrial biopsy were compared, the Pipelle sampler and the Tao brush.
Main outcome measures: Successful (informative) completion of the investigation, acceptability of the investigation method to women, women's satisfaction with clinic care in the short term and at 10 months and 2 years of follow-up, and cost-effectiveness to the end of investigation.
Results: Minor adverse events (e.g. shock, patient distress) did not occur for ultrasound, but occurred in 16% and 10% of women for hysteroscopy and biopsy procedures respectively. Pipelle biopsy provided an acceptable endometrial sample for 79% of moderate-risk women, but only 43% of high-risk women. The Tao brush gave similar performance in moderate-risk women (77%), but was more successful than the Pipelle sampler in postmenopausal (high-risk) women (72%). There were significantly more successful visualisations for ultrasound than for hysteroscopy in both the low-risk and the moderate-risk group, and a similar but non-significant trend in the high-risk group. Ultrasound was significantly better than hysteroscopy at detecting fibroids, but hysteroscopy significantly better for polyps. At the 10-month follow-up, high-risk women who had been investigated by hysteroscopy (with biopsy) had the most positive views of their clinic experience, but this effect had largely disappeared by 24 months. In the moderate-risk group, the subgroup randomised to biopsy alone gave the most negative responses about their clinic experience and health now. Women wishing they had more investigation comprised 22% of moderate-risk women and 38% of low-risk women, but only 14% of postmenopausal women. At follow-up the moderate-risk women (with menstrual bleeding problems), compared with postmenopausal women, had much worse ratings for clinic experience and health now. Resource use tended to be higher in the moderate- and low-risk women. There was minimal difference in cost-effectiveness between investigation options in the high-risk group, with the option involving hysteroscopy being marginally better than ultrasound. The most cost-effective investigation in the moderate-risk group was biopsy alone and in the low-risk group ultrasound.
Conclusions: Decision-making about investigation would be clarified if postmenopausal women were studied separately from premenopausal women with menstrual bleeding problems. For postmenopausal women exclusion of cancer is a main objective, so once investigation has been completed discharge follows, but in the woman with abnormal menstrual bleeding, even if serious pathology is excluded, the original presenting symptoms require management. About 60% of premenopausal women with abnormal bleeding reported that their symptoms were not 'much improved' at 10 months. Research is needed to understand this phenomenon, and to explore ways to

integrate patient factors into optimising evaluation and treatment. The significance of benign pathologies in this group also requires clarification. Given the relatively small differences observed in cost-effectiveness, there is justification for allowing other issues (such as clinician preferences and women's perspectives) to influence decisions as to the investigation method. There is scope to make better use of patient factors to inform decisions as to the most efficient and acceptable method of investigation for an individual woman. Additional analyses, using data available as a result of this study, will contribute to this agenda.

Contents

List of abbreviations

AUB	abnormal uterine bleeding
B	women randomised to both Tao brush and Pipelle biopsies
CAH	complex atypical hyperplasia
CER	comparative evidence required
CI	confidence interval
CONSORT	Consolidated Standards of Reporting Trials
CSSD	central sterile supplies department
CSW	clinical support worker
CVA	cerebrovascular accident
D&C	dilatation and curettage
DNA	did not attend
GA	general anaesthetic
GAPS	Gynaecological Audit Project in Scotland
GHQ	General Health Questionnaire
H	women randomised to hysteroscopy
HRT	hormone replacement therapy
ISD	Information and Statistics Division
ITT	intention to treat
LH-RH	luteinising hormone-releasing hormone
LMP	last menstrual period
N_t	number of women randomised to investigation (intention to treat)
N_m	number of women where investigation medically possible
N_c	number of women where investigation successfully completed
NA	not applicable
NEO FFI	NEO Five Factor Inventory
NF	nothing found
NPV	negative predictive value
O_u	number of women with outcome present
ODP	operating department practitioner
OMEE	outpatient methods of endometrial evaluation
OPD	outpatient department
OR	odds ratio
OR_{McN}	McNemar's odds ratio
P	women randomised to Pipelle biopsy
PPV	positive predictive value
PSSRU	Personal Social Services Research Unit
RIE	Royal Infirmary, Edinburgh
SIGN	Scottish Intercollegiate Guidelines Network
STARD	Standards for Reporting of Diagnostic Accuracy
T	women randomised to Tao brush biopsy
TVUS	transvaginal ultrasound
U	women randomised to transvaginal ultrasound

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices in which case the abbreviation is defined in the figure legend or at the end of the table.

Executive summary

Objectives

To compare three outpatient methods of endometrial evaluation in terms of performance, patient acceptability and cost-effectiveness.

Methods

Design

Pragmatic unblinded trial randomised separately within three groups determined by risk of endometrial cancer: high risk (postmenopausal women), moderate risk (premenopausal women either aged ≥40 years, or aged <40 years but with specific risk factors for endometrial cancer) and low risk (premenopausal women aged <40 years).

Setting

The gynaecology outpatient clinic of a large city hospital in Edinburgh, Scotland.

Subjects

Women referred for investigation and management of abnormal bleeding between January 1999 and May 2001 (n = 683 randomised).

Interventions

Investigations were: blind biopsy alone, hysteroscopy with biopsy, ultrasound evaluation including transvaginal ultrasound, and, in the low-risk group, the option of no investigation. To ensure adequate evaluation of all women, combinations of investigations were assigned, with the alternative options for a particular risk group as far as possible reflecting, at the time of funding application, consensus clinical practice for women with such risk. Within this design, two devices for obtaining endometrial biopsy were compared, the Pipelle sampler and the Tao brush.

Main outcome measures

Successful (informative) completion of the investigation, acceptability of the investigation method to women, women's satisfaction with clinic care in the short term and at 10 months and 2 years of follow-up, and cost-effectiveness to the end of the investigation.

Results

Overall 67% of those approached about the study were recruited. Recruitment met the target for postmenopausal women (n = 200, 100% of target) and nearly met it for moderate-risk women (n = 326, 82%), but was unsuccessful for low-risk women (n = 157, 52.3%), mainly because of changes in referral patterns and in investigation practice for this group. Over 90% of women completed all their recruitment questionnaires, 82% completed all their randomised investigations and over 83% returned their review of the clinic visit. There were high rates of follow-up to 10 months (77%) and of case-note review (98%).

Minor adverse events (e.g. shock, patient distress) did not occur for ultrasound, but occurred in 16% and 10% of women for hysteroscopy and biopsy procedures respectively. More women reported biopsy and hysteroscopy as markedly unpleasant, and for both these methods after-effects (bleeding and abdominal discomfort) were common. Nevertheless, the vast majority of women (87%) were reassured by their clinic visit and glad they had their investigation (94%), and overall 78% of women thought that their clinic visit was very or extremely worthwhile. There were only modest differences between investigations in these positive views. Women who had 'biopsy only' expressed greater wish to have had more investigation.

In high- and moderate-risk women, 15% intention to treat (ITT) of the Pipelle biopsies and 11% of hysteroscopies could not be undertaken for medical reasons, mainly failed insertion. Pipelle biopsy provided an acceptable endometrial sample for 79% of moderate-risk women, but only 43% of high-risk women. The Tao brush gave similar performance in moderate-risk women (77%), but was more successful than the Pipelle sampler in obtaining adequate samples in postmenopausal (high-risk) women (72%, $p < 0.001$). More women preferred the Tao brush than preferred the Pipelle sampler. Furthermore, adequate samples were more likely to be obtained if biopsy was undertaken at the time of hysteroscopy for both the Pipelle ($p = 0.027$) and Tao brush ($p = 0.002$).

There were significantly more successful visualisations for ultrasound than for hysteroscopy in both the low-risk (97% vs 65%, intention to treat, $p = 0.003$) and the moderate-risk group (88% vs 77%, $p = 0.002$), and a similar but non-significant trend in the high-risk group. Ultrasound was significantly better than hysteroscopy at detecting fibroids (32% vs 13%, $p = 0.006$), but hysteroscopy significantly better for polyps (13% vs 4%, $p < 0.001$).

At the 10-month follow-up, high-risk women who had been investigated by hysteroscopy (with biopsy) had the most positive views of their clinic experience, but this effect had largely disappeared by 24 months. In the moderate-risk group, the subgroup randomised to biopsy alone gave the most negative responses about their clinic experience and health now. Women wishing they had more investigation comprised 22% of moderate-risk women and 38% of low-risk women, but only 14% of postmenopausal women. At follow-up the moderate-risk women (with menstrual bleeding problems), compared with postmenopausal women, had much worse ratings for clinic experience and health now, in that less than half of them judged their symptoms 'much improved' by 10 months and one-quarter reported that their problem had not been cured.

Resource use tended to be higher in the moderate- and low-risk women, because of the need to manage their abnormal bleeding symptoms. There was minimal difference in cost-effectiveness between investigation options in the high-risk group, with the option involving hysteroscopy being marginally better than ultrasound (£88/woman). The most cost-effective investigation in the moderate-risk group was biopsy alone (saving £128–212/woman compared with the other options) and in the low-risk group ultrasound (£74–452/woman better).

Conclusions

This study has highlighted the complexity of the investigation pathways travelled by women referred for abnormal bleeding. Decision-making about investigation and understanding would be clarified if postmenopausal women were studied separately from premenopausal women with menstrual bleeding problems. For postmenopausal women exclusion of cancer is a main objective, so once investigation has been completed discharge follows, but in the woman with abnormal menstrual bleeding, even if serious pathology is excluded, the original presenting symptoms require management.

About 60% of premenopausal women with abnormal bleeding reported that their symptoms were not 'much improved' at 10 months. Research is needed to understand this phenomenon, and to explore ways to integrate patient factors into optimising evaluation and treatment in these cases. The significance of benign pathologies in this group also requires clarification.

Given the relatively small differences observed in cost-effectiveness, there is justification for allowing other issues (such as clinician preferences and women's perspectives) to influence decisions as to the investigation method. The clinicians expressed interest in the Tao brush being made available for their use. Its introduction would have resource implications, in particular the training of pathology staff. The Tao brush is superior in obtaining adequate samples, so it should be considered the method of choice for postmenopausal women, or at least be readily available as a back-up technique where Pipelle sampling has failed.

At the time of investigation ultrasound was much more acceptable to women than hysteroscopy and biopsy, but hysteroscopy was not more unpleasant to women than biopsy. Women having hysteroscopy were pleased to have had the investigation and women having this randomisation option were least likely to have wanted more investigation, whereas those having biopsy only wished that they had had more investigation.

There is scope to make better use of patient factors to inform decisions as to the most efficient and acceptable method of investigation for an individual woman. Additional analyses, using data available as a result of this study, will contribute to this agenda.

Chapter 1
Introduction

Abnormal uterine bleeding (AUB) is an important symptom of both benign and serious gynaecological disease. AUB is the single most common reason for gynaecological referral.[1] In particular, postmenopausal bleeding may be an early symptom of endometrial carcinoma. Excessive menstrual blood loss affects 10–30% of menstruating women and in the order of 70% of all gynaecological consultations in the perimenopause and postmenopause.[2] Abnormal perimenopausal or postmenopausal bleeding is associated with endometrial carcinoma in approximately 10% of cases.[3,4] In 18–40% women benign focal lesions, such as endometrial polyps and fibroids, are common.[5] With regard to endometrial carcinoma, the significance of abnormal bleeding depends quite strikingly on demographic factors. If there is postmenopausal bleeding endometrial carcinoma is of particular concern. Abnormal bleeding in a premenopausal woman is, however, not a symptom of immediate concern with respect to cancer, because there are a number of potential physiological explanations, and also because the incidence of uterine cancer under 40 years of age is very low.[6,7] It has been calculated that 3000–4000 women under 40 years with abnormal bleeding would have to be evaluated to detect one case of endometrial cancer.[6] This has implications for decisions regarding evaluation of AUB.

Altough the most important reason for endometrial evaluation remains exclusion of serious pathology (e.g. endometrial cancer), evaluation of the uterine cavity is also undertaken in cases of excessive monthly menstrual blood loss (menorrhagia) and when irregular vaginal bleeding is reported. In the1996 Gynaecological Audit Project in Scotland (GAPS) report,[8] the proportions of endometrial assessments for these indications were: postmenopausal bleeding 21%, menorrhagia 27% and irregular bleeding 28%. Thus, a further factor that should impact on decisions about undertaking investigation is the nature of the symptoms of AUB. A small amount of postmenopausal bleeding may be very little trouble to the woman, so it is often the case that if a selected investigation excludes endometrial carcinoma then there is no need for further management of the minimal symptoms. In the younger premenopausal woman, however, AUB is more likely to be troublesome. Therefore, regardless of any wish to exclude serious pathology, there is also a requirement to investigate so as to diagnose the aetiology of troublesome symptoms to inform appropriate therapeutic management. The GAPS project and subsequent GAPS reaudit demonstrated that 86% of investigations conducted to evaluate the endometrium were performed appropriately: 77% were conducted in women over 40 years of age with AUB; 2% were performed in women under 40 years with AUB following a failure of medical management.[9,10]

Currently the most commonly used technologies for outpatient evaluation of the endometrium are biopsy, hysteroscopy and transvaginal ultrasound (TVUS). The choice of which modality to use for investigation will be influenced by: risk of endometrial disease, menopausal status, local availability of investigative options and whether exogenous hormones are being taken. There is evidence that endometrial sampling alone may miss lesions in between 10 and 33% of cases.[11] Thus, the use of complementary methods of endometrial assessment, hysteroscopy and TVUS, have been described and incorporated into assessment.[12]

Techniques for outpatient evaluation of the endometrium

Endometrial biopsy

Endometrial sampling provides tissue for histological diagnosis and thus exclusion of premalignant or malignant disease. Dilatation and curettage (D&C) has been the traditional gold-standard investigation for AUB. Alternative methods have been sought as a D&C requires general anaesthesia and is a blind procedure for tissue sampling. Two-thirds of women examined by this method have been reported to have less than half the uterine cavity sampled and 16% less than one-quarter.[13] The main disadvantage of a blind endometrial biopsy is that focal lesions such as polyps, submucous fibroids and localised pathological lesions may be missed.

Pipelle sampling

The most common method for collection of an endometrial biopsy in recent years has been use of the Pipelle endometrial sampler (Eurosurgical, UK). Use of this sampling device was the first line method for 66% of Scottish consultants in 1996.[8] The Pipelle endometrial sampler obtains by aspiration a small sample of endometrium for histological assessment. Importantly, it may be performed without dilatation of the cervix and consequently there is usually minimal discomfort associated with the procedure, which may be conducted in an outpatient setting. The main advantages of this device are familiarity in use, low cost, routine pathological assessment and its suitability for the outpatient environment. It remains, however, an invasive procedure and for some women is associated with discomfort. In postmenopausal women with an atrophic endometrium it is not unusual for the Pipelle sampler to fail to obtain a sample, but this circumstance can nevertheless be reassuring to the clinician, suggesting that no serious pathology, likely to be more productive of histological tissue, is present. Regardless of whether or not a sample is obtained, concerns have been expressed that only a very small area of the endometrium is sampled with this device. An assessment of the biopsies so obtained did, however, agree with the posthysterectomy diagnosis in 84% of cases.[14]

Only very recently have systematic reviews with comparison of outpatient sampling devices been available in the literature. Clark and colleagues[15–17] have reported that outpatient endometrial sampling is a successful procedure and, when adequate samples are obtained, has a high overall accuracy in diagnosis of endometrial malignancy. Although rigorous criteria were used for the assessment of studies included in the review, it was still unable to compare individual sampling techniques with reliability. Furthermore, difficulties may arise in the assessment of postmenopausal women with AUB as collection of a sufficient sample for histological analysis is difficult in those women with endometrial atrophy. It is recognised that outpatient sampling in women with AUB will fail to collect a sample adequate for histological assessment in up to 10% of women.[18]

Tao brush sampling

A newer development for cytological assessment of the endometrium is the Tao brush sampler (Cook, UK), which permits sampling of the surface of the endometrium without excessive manipulation and produces a sample uncontaminated by material from the lower genital tract.[19] The brush technique has the advantage that microbiopsies are also produced, especially in cases of endometrial hyperplasia and carcinoma, thereby allowing quality control of the cytological diagnosis. In 79.9% of cases tissue fragments, suitable for histological assessment, are obtained.[19] The Tao brush is inserted into the uterine cavity with a plastic sheath covering the brush. Once inserted, the sheath is withdrawn to expose the brush. The brush is then rotated through 360 degrees clockwise and then anticlockwise to obtain the sample. The brush is resheathed before being withdrawn from the uterine cavity and bottle-brushed into a collection tube containing a cytofixative. It is possible that the brush sampling technique may be more acceptable to women than the aspiration sampling technique, especially if the uterine cavity is atrophic. The Tao brush, like the Pipelle sampler, is suitable for outpatient use but remains an invasive procedure. The Tao brush incurs slightly greater costs, as a non-standard pathological assessment procedure is required for analysis of samples collected by this technique. Specific training is necessary for pathologists to analyse samples collected in this manner. Hence, a rigorous health economic assessment is required before wider implementation of use of this method for endometrial biopsy assessment. Importantly, this technique of tissue collection does not interfere with subsequent histological studies of endometrium sampled by conventional curettage or aspiration (Pipelle).

Hysteroscopy

Hysteroscopy permits endoscopic evaluation of the uterine cavity, with video-recording also possible (e.g. if there was need for a second opinion or as a valuable aid for technique instruction). Hysteroscopy can be undertaken in the outpatient setting without analgesia. It is reported to be well tolerated.[20] In Scotland in 1996 it was, however, most often undertaken as a theatre procedure, under general anaesthesia. The GAPS audit[8] reported that in Scotland hysteroscopy was six times more common as an inpatient than an outpatient procedure. In 100 consecutive procedures conducted at the Royal Infirmary, Edinburgh (RIE), in 1997,[10] 29% of procedures involved a hysteroscopy, of which slightly less than one-third (28%) were conducted in an outpatient environment. Direct visualisation of the uterine cavity with this technique allows diagnosis not only of a visible carcinoma, but also of polyps and of fibroids within the uterine cavity. A visually suspected endometrial carcinoma would require a complementary histological diagnosis. In cases where no intrauterine pathology is visible it is

usual practice to complement hysteroscopy with an endometrial biopsy, to exclude early intrauterine pathology of a premalignant or malignant nature.

Advantages of hysteroscopy are thus claimed to be that the biopsy is no longer blind and so successful sampling is more likely, that the view of the intrauterine cavity permits identification of other benign disease, for example, polyps or fibroids, and that, where appropriate, therapeutic intervention may be possible, since it is possible to remove small intrauterine polyps at the time of initial hysteroscopy.

Furthermore, the woman may be provided with reassurance. She may have chosen to visualise the procedure via the video-screen facility (only possible with a conscious patient, usually in an outpatient setting). Disadvantages of the use of hysteroscopy in assessment of AUB include the cost of purchase and maintenance of the outpatient equipment, familiarity with the technique among gynaecologists, and the possibility that abnormalities of no clinical significance are seen and therefore treated. Hysteroscopy has increasingly replaced the traditional D&C for the evaluation of AUB.[21]

Pelvic ultrasound, including transvaginal ultrasound

Pelvic ultrasound is the least invasive of the outpatient techniques that can be used for visualisation of the structure of the uterus and for visualisation of the thickness of the endometrium. The method of pelvic ultrasound that best visualises the endometrium is TVUS. This in the postmenopausal woman allows accurate measurement of endometrial thickness. Use of TVUS for measurement of endometrial thickness has a high diagnostic accuracy in the identification of endometrial cancer.[22] A meta-analysis of postmenopausal women reported that 96% of women with endometrial malignancy and 92% of women with any endometrial disease had a measured endometrial thickness greater than 5 mm.[4] In the context of unscheduled bleeding on hormone replacement therapy (HRT), ultrasound has a reduced diagnostic specificity. A subsequent meta-analysis (pooling data from four studies) demonstrated that a negative ultrasound result of 5 mm or less reduced the risk of disease by 84%. It has been pointed out that the pretesting risk of disease will determine whether the above finding is sufficient to rule out disease. Moreover, data derived from the four studies had a wide 95% confidence interval (CI) for reduction in risk (54–94%) with a negative result. Thus, a range of factors requires consideration when determining best use of TVUS and the optimum threshold for endometrial thickness.[23,24]

In contrast, in the premenopausal woman the role for pelvic ultrasound is in the diagnosis of structural abnormalities as an explanation for AUB, such as submucous fibroids or endometrial polyps. The advantages of pelvic ultrasound are that it is non-invasive, there is no need for anaesthesia and, in certain cases, it provides additional information that may assist both diagnosis and treatment choice. In clinical practice it is common for radiologists and radiographers to undertake ultrasound investigation by both methods, starting with abdominal, which requires a full bladder to provide optimal visualisation of the ovaries and uterine structure, and then after the woman has emptied her bladder, proceeding to TVUS, which provides the best view of the endometrium. In an obese woman a transabdominal pelvic ultrasound may be less than optimal.

At the time of the GAPS audit in 1996[8] only one-quarter of Scottish consultants agreed that ultrasound was an appropriate first line approach for endometrial evaluation for postmenopausal women. In this audit women were selected on the basis of having had one of the other methods of endometrial evaluation, but a record was made of whether the woman also had an evaluation by ultrasound. In 1995 only 3.5% of the sample had, and by 1997 this had increased to only 7%.[8,9] This report of low combined use of ultrasound with one of the other methods of endometrial evaluation may mask a high and successful (sole) use as a first line method, but in view of consultant attitude at this time, reported above, this is unlikely.

The usefulness of ultrasound may be limited by interobserver error and the variable ability to exclude polyps and fibroids. A technique known as saline infusion sonohysterography, involving the introduction of a small volume of saline into the uterine cavity, may enhance the sonographic view and aid the diagnosis of intrauterine lesions.[25] The sensitivity of saline infusion sonohysterography in the identification of endometrial polyps is reported as 83% compared with 16.7% with TVUS.[26,27]

Clinical decisions regarding method of endometrial evaluation

At the time of submission for funding there were no published randomised trials that compared

outpatient methods of uterine evaluation (biopsy, hysteroscopy, transvaginal ultrasound) for the diagnosis of endometrial abnormality. Even in 2002 there were, to the author's knowledge, no published studies assessing the diagnostic accuracy of hysteroscopy, endometrial biopsy and ultrasound that satisfied the criteria for optimal study design.[2] These authors draw attention to the fact that validity of diagnostic studies will reflect the clinical context in which the assessment is undertaken. Thus, to date, some of the best published evidence for endometrial evaluation method is for ultrasound. In a series of premenopausal and postmenopausal women with abnormal bleeding, where there was an abnormality prevalence of 42% (carcinoma 2%), initial evaluation by ultrasound would have identified two subgroups: one with abnormal ultrasound scan with a probability of abnormality of 87%, and the other, with normal ultrasound, with a probability of only 3% of any abnormality. This could have reduced the number of hysteroscopies required by 50%.[28]

It is therefore surprising that ultrasound is so little favoured by gynaecologists (e.g. see Refs 29 and 30). The use of hysteroscopy is growing, and this method is being predicted to become "as routine in the twenty-first century as D&C has been in the twentieth".[31] This is a rather dubious recommendation, given the widespread inappropriate use of D&C, despite the initial lack of evidence as to its efficacy, and despite growing evidence as to its inefficacy in many cases. There is a worrying similarity in the history of hysteroscopy, which is rapidly diffusing into clinical management of abnormal bleeding, despite the lack of published evidence of its benefit over blind biopsy.[30] The favouring of hysteroscopy over ultrasound may well be a reflection of the fact that hysteroscopy gives immediate visualisation of the uterine cavity, and of the fact that it is a gynaecological procedure undertaken in person by the clinician managing the patient. More recently, quantitative systematic reviews have been undertaken to evaluate the individual value of outpatient hysteroscopy,[17] endometrial biopsy[15] and ultrasound[4,23] for the detection of endometrial hyperplasia and cancer. The dearth of information concerning optimal investigation strategy with regard to effectiveness and cost-effectiveness has been identified.[2]

TABLE 1 *Evidence on outpatient methods of endometrial evaluation, as at 1996/97*

	Blind biopsy		**Hysteroscopy + biopsy**	**Ultrasound**
	Pipelle	**Brush**		
Diagnostic success				
Achieving evaluation				
Cannulation	96%	96%	89–96%	NA
Sample obtained	75–80%	95%	CER	NA
Overall	74%	91%	CER	90–97%
Detecting endometrial carcinoma				
Sensitivity	45–84%	CER	CER	96%
Specificity	99%	99%	99%	60–68%
Other endometrial disease			In uterine cavity only	In cavity and in uterine structure
Sensitivity	0%	0%	CER	CER
Clinical relevance	NA	NA	CER	CER
Adverse events	Perforation Infection	Perforation Infection	Perforation Infection Vasovagal episode	None
Patient acceptability	CER	CER	CER	CER
Clinician attitudes	Favourable	CER	Increasingly favourable	Cautious
Procedure cost (notional)				
Clinic appointment and evaluation	£40	£40	£40	£40
Pathology and device costs	£29	£40	£29/£40	£0
Total	£69	£80	£69/£80	£40

CER, comparative evidence required; NA, not applicable.

Evidence to inform choice of endometrial evaluation modality

The evidence available at the time of submission for funding is summarised in *Table 1* and brought up to date in the following text.

Diagnostic performance

Success at detection of endometrial pathology (sensitivity)

Concern has been expressed that the Pipelle sampler only samples 4% of the endometrial cavity[14] and that consequently focal cancers may be missed. However, where samples are successfully obtained the method has reported sensitivity of 44.6–84%.[8,32] A recent review of 33 reports summarising 13,598 D&Cs and 5851 office hysteroscopy procedures demonstrated that D&C had a greater complication rate, but that the diagnostic accuracies of the two techniques were comparable.[33] The sensitivity of the technique has been shown to improve for all types of endometrial disease if Pipelle sampling is complemented with ultrasound.[8,34] The main disadvantage of use of the Pipelle is the relatively high proportion of inadequate samples, particularly so among postmenopausal women (the group where unscheduled uterine bleeding is a strong predictor of endometrial malignancy).

Inadequate tissue sampling occurs in up to 68% of postmenopausal women and 21.5% of premenopausal women. The success of this biopsy technique is thus dependent on the adequacy of the tissue sample obtained. The reported overall percentage of failed insertions or inadequate samples for all biopsy techniques using a suction sampling technique (Pipelle, Gynoscann, Accurette, Novak curette, Vabra aspirator and Z sampler) was 7% (95% CI 5–8%). The Pipelle had a failure rate of 8% (95% CI 6–11%.[15] The percentage of specimens inadequate for histological analysis was reported as 15% of biopsies overall (95% CI 12–17%) and among 13% of Pipelle samples (95% CI 10–16%). Sample inadequacy was greatest among postmenopausal women (22%, 95% CI 18–26%.[15] In the GAPS II audit[9] 5% of inpatient theatre evaluations in 1995 were as a result of a failed outpatient endometrial biopsy, and in 1997 this number had increased to 7%, despite no absolute change in the overall proportion of outpatient procedures.

Thus, data derived from the present study have allowed exploration of predictive modelling of the characteristics of women and evaluation methods where there are likely to be technical difficulties with an outpatient evaluation.

The Tao brush endometrial sampling device achieves a higher rate of successful samples than the Pipelle, 95% compared with 80%.[19] Where other samples are obtained the accuracy of diagnosis is good. In a series of 656 hysteroscopy specimens Maksem and Knesel[19] had no inadequate samples. Diagnostic accuracy was 100% for atypical hyperplasia and carcinoma, and 92.5% overall. Bistoletti and Hjerpe[35] reported routine use of endometrial cytology (Tao brush) in clinical practice and were able to identify 97% of cases of atypical hyperplasia and 96% endometrial carcinoma.[36]

Ultrasound can also fail to obtain a measurement (3–10%).[28,37] Where measurement is made it has been found to have a very high sensitivity to abnormal histological endometrial findings (93% to 96%).[37,38] Both studies have suggested that TVUS could be used to select the women who would benefit from endometrial biopsy, and that with 4 mm as a cut-off limit, TVUS of the endometrium could exclude endometrial abnormality with reasonable certainty. Assessment of the ultrasonic morphology of thickened endometrium may be of some value in differentiating between hyperplasia, endometrial polyps and endometrial carcinoma.[39] TVUS is an accurate method for measuring the depth of invasion in women with endometrial carcinoma.[37]

Hysteroscopy has a modest failure rate, when it is not possible to pass the hysteroscope along the cervical canal into the uterine cavity. In a series of 2500 outpatient procedures, hysteroscopy was not possible for 4% of women and was incomplete for 7%.[31] In such cases it will normally also be unlikely that an endometrial biopsy is possible.

Hysterscopy has been reported to be well tolerated by women and importantly has enabled detection of endometrial lesions following a negative endometrial sample or D&C. Shushan and colleagues[40] and Revel and Shushan[41] reported that following normal endometrial histology of a 'blind' sample from the uterine cavity, hysteroscopy permitted a directed biopsy of suspicious intrauterine lesions which in several cases exposed a diagnosis of endometrial cancer. The study by Gimpelson and Rappold[1] that compared hysteroscopically directed biopsy and D&C reported that the result of curettage was in agreement in 80.8% of cases. Hysteroscopy exposed more information than curettage in 16% of women, whereas curettage revealed more than

hysteroscopy in only 3.3% of women. More recently, a prospective study by Gebauer and colleagues[42] has demonstrated that curettage alone is an inadequate procedure for the detection and removal of endometrial polyps. Curettage alone detected polyps in 43% of women known to have had an endometrial polyp detected at hysteroscopy. A combination of ultrasound and outpatient hysteroscopy plus endometrial biopsy has been reported to have efficiency equal to inpatient hysteroscopy and curettage for the investigation of AUB.[43]

Thus, invasive assessment of the uterine cavity is only justified in the evaluation of AUB to detect either benign endometrial pathology or endometrial cancer. Hysteroscopy has recently been reported to have a sensitivity of 94.2%.[44] However, if hysteroscopy is used alone, without complementary biopsy, it has a sensitivity of only 58.8%.[45]

Although Pipelle and brush samples of the endometrium allow diagnosis of both cancer and premalignant change, they do not identify other endometrial disease such as intrauterine polyps and fibroids. In premenopausal women the main role of evaluation is the diagnosis or exclusion of structural abnormalities as a cause for AUB. Ultrasound has been shown to be effective in predicting the presence or absence of endometrial pathology as determined by hysteroscopy and curettage. In one study 99% of submucous fibroids and 89% of endometrial polyps were detected by scanning.[46] In a further study evaluating 97 women with menorrhagia ultrasound correctly detected the presence of benign uterine enlargement, the number and size and location of fibroids, submucous polyps and bicornuate uteri which had not been detectable clinically. The ultrasound diagnosis increased the accuracy of the clinical diagnosis and assisted in the proper choice of treatment.[47] There is no published comparative evidence of the effect of hysteroscopy on diagnosis of endometrial disease other than cancer, although anecdotally it is believed to be superior.

Success at avoidance of erroneous diagnosis of disease (specificity)

With regard to endometrial carcinoma, biopsy by either Pipelle or brush has very high specificity, regardless of whether a hysteroscopy has been performed, that is, false-positive diagnoses are extremely rare. TVUS does not detect malignancy directly, but an endometrial thickness greater than 3 mm has been found to include virtually all cases of carcinoma. For screening women who report postmenopausal bleeding different cut-offs may be applied to endometrial thickness measurements. The lower the cut-off used, the more sensitive and less specific the test. In practice, many clinical units use greater than 4 mm, but recent guidelines recommended greater than 3 mm.[24] In the absence of cancer, high endometrial thickness measurements have been found to indicate uterine polyps, so there is a potential for diagnostic advance if the follow-up biopsy is conducted with a complementary hysteroscopy. Indeed, since the incidence of focal lesions (e.g. polyps and fibroids) in women who describe AUB has been reported to lie between 46 and 74%,[31,48–50] there has been an argument to favour hysteroscopy and endometrial biopsy as the most cost-effective approach in early evaluation of AUB.[41]

The endometrial evaluation methods ultrasound and hysteroscopy can also detect structural abnormalities of unknown effect, and some benign endometrial disease. In another series[28] which made more specific assessment than solely endometrial thickness, the false-positive rate was only 10%. The value of this depends on the extent to which such abnormality needs to be treated (i.e. is the cause of the complaint) and may be treated. For example, in a woman of 45 with excessively heavy periods that do not respond adequately to medical treatment, and who is amenable to hysterectomy, the treatment offered is unlikely to depend on a morphological characterisation of the endometrium. Where an evaluation technique provides information of uncertain diagnostic or therapeutic value then it could be said to have low specificity in respect of clinically relevant findings. A surrogate measure of relevance (specificity) of such endometrial findings will be the extent to which therapeutic management is expedited.

Adverse events

Outpatient procedures do not require anaesthetic, so the well-documented risks of general anaesthesia are avoided. Potential adverse events of an endometrial biopsy are the theoretical introduction of infection and, rarely, uterine perforation. Adverse events associated with hysteroscopy would again include the theoretical risk of introduction of infection, risk of uterine perforation and very occasionally an episode of cervical shock may occur (vasovagal episode). There are no documented adverse events associated with TVUS.

Cost-effectiveness of evaluation procedures

In 1997, according to the Edinburgh Royal Infirmary Finance Department, there was little

difference in notional cost between the three methods of outpatient evaluation (as summarised in *Table 1*), contrary to what would have been expected, considering the differences in time taken to complete the investigations, staffing configurations used, and the range of cost of facilities, equipment and disposables involved. However, as outpatient evaluations become more common and the requirement for practitioners to identify and utilise the most cost-effective options becomes more widespread there will need to be more careful consideration of relative costs and associated outcomes. When choosing a method of endometrial evaluation it is not just the cost of the procedure that should be considered, but also the full cost-effectiveness of the evaluation procedure relative to other available methods.[51]

There has been growing realisation of the importance of an agreed way of managing the assessment and introduction of new technologies in healthcare, as otherwise the likely consequence of local decisions about implementation is inequity of provision of health care services. In this study the methodological approach used in assessment of the cost-effectiveness of endometrial evaluation technique is that suggested by the White Paper *Designed to care*.[52]

Patient acceptability

Acceptability of an evaluation method, to the woman being investigated, will be a composite of discomfort or distress, or both, while undergoing the procedure, quality of reassurance, whether she is concerned about sinister disease, duration of anxiety, convenience, avoidance of the medical risk and general anaesthetic that pertained to some alternative, perhaps more invasive technique, and the extent to which the evaluation is perceived to expedite or optimise treatment. Where there is little difference in sensitivity in terms of diagnosing serious disease, patient acceptability is an important factor in decision-making with regard to evaluation options. Few published studies of hysteroscopy have formally assessed patient acceptability[35] and at the time of funding none was found in the literature.

Gynaecologist attitudes

There is considerable discrepancy between consultants' expressed beliefs and their practice. For example, two-thirds of consultants agree with the (cost-based) principle that women under 40 do not require an endometrial evaluation, but only one-third adopt this principle in practice.[8,9] Where evaluations are unnecessary it is of interest to elucidate factors that contribute to their being undertaken. For example, to what extent is it a clinical response to a very anxious woman? While 80% of Scottish gynaecologists agreed with the principle that outpatient sampling procedures are the endometrial evaluation techniques of choice, in practice only half of sampling procedures reviewed were undertaken in the outpatient setting.[9] Thus, is it clinician or woman's characteristics that determine the number of theatre (inpatient) procedures undertaken? The GAPS audit[8] found that 11% of Scottish consultants used D&C under anaesthetic as their first line method of evaluation. Their main reason was given as: "not convinced of adequacy of endometrial biopsy".

Overview

Choice of evaluation modality thus requires a rigorous comparison between the three endometrial investigation techniques as used in usual clinical practice, in terms of:

- diagnostic performance
- patient acceptability and reassurance
- adverse events
- cost-effectiveness
- gynaecologist attitudes.

Chapter 2
Study design

Aims

The study aimed to provide evidence of acceptability to women, adequacy and cost-effectiveness of outpatient evaluation of the endometrium, to demonstrate feasibility of evaluation in the outpatient clinic, and to familiarise clinicians with the execution of newer methods of outpatient evaluation, and the use of the results of such diagnostic evaluation.

Objectives

The main objectives of the study were:

- to compare the three outpatient methods of diagnosis of endometrial abnormality (ultrasound, blind biopsy and hysteroscopy plus biopsy) in terms of:
  - diagnostic performance (i.e. completion rate, time taken, success at obtaining diagnostic material/view, abnormalities detected, accuracy, after-effects)
  - adverse events
  - women's views (immediate) on experience of clinic and investigation(s)
  - women's views (longer term, at 10 and 24 months postrecruitment) on clinic investigation(s) and self-report of outcome and health
- to compare Tao brush endometrial sampling with Pipelle biopsy in terms of diagnostic performance, and women's reports and expressed preferences
- to perform a cost evaluation and comparison of cost-effectiveness of the three outpatient methods of endometrial investigation and, within biopsy, of Tao brush compared with Pipelle sampler.

Secondary objectives were:
- to elucidate the characteristics of women for whom outpatient evaluations are unsuccessful owing to technical difficulties, or are unpleasant
- to describe the pattern of adjunctive evaluations (outpatient and inpatient) undertaken and resource use across the duration of the study, and examine its association with a woman's characteristics and secular trends
- to examine clinician attitudes to methods of endometrial evaluation over the course of the study.

Design

The design chosen was a randomised controlled trial, with the investigations randomly allocated being perforce unblinded. Since the potential explanation for the abnormal bleeding [primarily whether it is serious disease (cancer) or not], and hence by association the woman's age and menopausal status, have such a strong influence on clinical decision-making, it is important that these factors are taken into account in the design. The study was therefore conducted within three groups, the groups reflecting the reproductive status and hence the significance of signs and symptoms, that is, the background risk of non-benign abnormality:

- high risk: postmenopausal
- moderate risk: premenopausal and either aged 40 years or over, or younger but with specific risk factors (polycystic ovarian syndrome, prior use of unopposed oestrogens or tamoxifen, obesity, diabetes or family history of endometrial cancer[8])
- low risk: premenopausal, age under 40 years and without specific risk factors.

It was considered important that the comparison of evaluation methods was undertaken in a setting as close as possible to normal clinic operation. For this reason, and to maximise clinician compliance with the study, a pragmatic design was used. After execution of the randomly assigned investigations the clinician could continue management of the patient unconstrained by the study, so that if further outpatient or inpatient investigations were indicated they could be offered in the normal way. For assigned ultrasound investigations the transvaginal method would be used wherever possible, but the investigation would be limited to abdominal if that was preferable for a particular woman.

As technology for evaluation of AUB develops the contending modalities become more and more similar in terms of standard diagnostic performance criteria. Increasingly, diagnostic evaluation choice

will be made on the basis of women's preferences or cost, or both. If cost is to be considered then it is important that this is realistic and longer term, including not just the cost of the evaluation utilised, but also costs of any additional clinic visits and evaluations that may be necessary, of suboptimal treatment choice and of disruption to a woman's life, and so on. Therefore a prospective study was required, with patient follow-up at 10 months and, for as many as possible, also at 24 months.

Ethical approval was granted in August 1997 by Lothian Research Ethics Committee (reference 1702/97/6/34).

Interventions

Description of interventions

All three 'interventions' were outpatient investigations and two of the three involved a biopsy. Implanted within the overall comparison of the three interventions was a comparison of two devices for obtaining pathological information:

- standard Pipelle endometrial biopsy sampler
- the newer Tao brush sampler.

The three interventions were:

- endometrial evaluation by blind biopsy, using the Pipelle sampler and/or Tao brush
- TVUS, usually in conjunction with abdominal ultrasound and in some cases substituted by abdominal ultrasound
- hysteroscopy with biopsy, using the Pipelle sampler and/or Tao brush.

The endometrial biopsy samplers and the other methods of endometrial evaluation have been described in detail in Chapter 1.

Assignment of interventions within groups

For ethical reasons, and to maintain clinical confidence in the study, women were randomised within the risk groups to sets of evaluation options chosen to reflect usual clinical practice with regard to precautionary investigation. The randomisation options therefore differed for each risk group, as follows (see also *Table 2*).

- High-risk group (postmenopausal): for these women there was, on the whole, consensus that symptoms of bleeding must be investigated thoroughly, so all women should receive both biopsy and visualisation (TVUS or hysteroscopy). Randomisation was therefore to one of two options: hysteroscopy with biopsy, or blind biopsy plus ultrasound.
- Moderate-risk group (premenopausal, and either aged ≥ 40 years, or < 40 years but with specific risk factors): it is this group for which there is most doubt as to optimum diagnostic strategy, but as a precaution every woman should receive at least a biopsy. The two visualisation procedures (ultrasound or hysteroscopy) were randomly assigned in addition to biopsy, in a factorial design, to ensure efficient statistical comparison of the

TABLE 2 *Randomisation assignments by risk group*

	Group		
	High-risk	**Moderate-risk**	**Low-risk**
Minimum investigation required	Pipelle biopsy and visualisation	Pipelle biopsy	Investigation not obligatory
Investigation options			
Biopsy[a]	All receive	All receive	Randomised to one or none of the three options
Ultrasound / Hysteroscopy (with biopsy)[b]	Randomised to one or other visualisation	Randomised factorially to either, neither or both 'visualisations'	
Number of randomisation subgroups	2 equal	4 equal	4 unequal (none 40%, other 3 options 20% each)

[a] The groups with high and moderate risk were biopsied using both Pipelle and Tao brush, whereas low-risk women randomised to biopsy were sampled using either sampler (50% each).
[b] If the woman was not randomised to hysteroscopy then the biopsy was blind. If the woman was randomised to hysteroscopy then the biopsy was undertaken at the same time (i.e. after visualisation).

procedures for number of women studied. Therefore, one-quarter of the women received neither ultrasound nor hysteroscopy, one-quarter of women received both, one-quarter received only ultrasound and the remaining quarter received only hysteroscopy.
- Low-risk group (< 40 years and without specific risk factors): there is consensus that in this age group investigation is not required on grounds of likely pathology. Just less than half of the women (40%) therefore underwent no investigation, receiving instead clinic management by medical treatment and surveillance. The remainder were offered one of the three interventions.
- Biopsy method: with regard to the biopsy component of the interventions, since the brush is as yet unproven it was ethically necessary that all women in the high- and moderate-risk groups underwent at least a Pipelle biopsy. The most powerful comparison of Pipelle biopsy with brush sampler is if all of these women have biopsy by both methods. However, in the low-risk group, for whom investigation may not be essential, women underwent only one biopsy, with half of the biopsies undertaken by Tao brush and half by Pipelle sampler.

Participants

The study included all women referred to the gynaecology outpatient clinic at Royal Infirmary Edinburgh, Scotland, for abnormal bleeding, but only if the managing clinician consented to the woman being approached about the study and the referral complaint of abnormal bleeding had been verified by that clinician. The study included women using oral contraception or HRT, but excluded those who were pregnant or who had difficulty reading or writing English. Subsequently, five women were recruited from satellite outpatient clinics.

Study measures

Data collection forms, reports and questionnaires

The scope of the study, and the exploratory nature of aspects of it, mean that a wide range of data was collected. Study variables were derived from clinician-completed forms and patient-completed questionnaires, and covered the time from initial appointment (recruitment) to 2 years later. The nature and timing of the forms, reports and questionnaires are summarised in *Table 3*.

Clinician-completed forms

The first paperwork completed by the managing clinician was the Consent Slip, confirming clinician agreement to the woman, who had already consented to the recruitment staff, being entered into the study and randomised, and also confirming the woman's risk group. Up to six further forms, depending on the investigations carried out, were completed by clinicians or the pathologist to record clinical information about the woman, the conduct of the study investigations and the findings. The clinical details form recorded presenting complaint, symptoms and main health problems, including contraception, hormonal treatment, date of last menstrual period (LMP) and any previous gynaecological problems. It also

TABLE 3 *Study forms, reports and questionnaires*

Timing	Forms completed by clinicians/pathologists	Patient-completed reports and questionnaires
Recruitment	Clinician recruitment slip Clinical details-of-patient form	Health questionnaire NEO Five Factor Inventory (FFI) General Health Questionnaire (GHQ) Clinic visit report
Assigned investigation performance form	Biopsy form Hysteroscopy (+ biopsy) form Ultrasound form	Biopsy report Hysteroscopy report Ultrasound report
Pathology results forms	Pipelle results Tao brush results	
Day after last investigation completed		Review of clinic attendance
Follow-up questionnaires		10-month follow-up 24-month follow-up

recorded clinician assessment of risk factors and risk group, and the clinician's preferences for investigation for this woman. For each investigation a performance form was devised, and each biopsy method had a results form designed specifically for the study. These forms are provided in Appendices 1 and 2. They were developed in collaboration with specialists in pathology [A Williams (AW)], transvaginal ultrasound [S Chambers (SC)] and gynaecology/hysteroscopy [HOD Critchley (HODC)], to deliver data necessary to achieve the study aims.

Forms completed by the woman around the time of investigation

The first stage in development of the forms for patient completion was to establish focus groups with women who had consulted for abnormal bleeding, to elucidate the issues from the patients' viewpoint. The methods and summarised findings for the focus groups are summarised in Appendix 3. Psychological theory together with focus group findings informed the development of patient questionnaires and reports, and the selection of established psychiatric health and personality scales to be used. In-house patient-completed reports and questionnaires are provided in Appendix 4.

At recruitment each woman was assessed in terms of a number of potential explanatory factors. This was so that her subsequent management, experience of investigation and ease of reassurance could be analysed in terms of these potential explanatory factors. The 'Health' questionnaire that was devised for the study assessed the woman's self-reported health, including surveillance of and coping with bodily symptoms, any ideas as to explanations for her abnormal bleeding, and (prior) attitudes to the various endometrial investigations.

Personality was assessed by means of Costa and McCrae's NEO Five Factor Inventory Form S[53] (NEO-FFI; NEO), which measures five major dimensions of personality. The NEO was developed by factor analytical methods integrated with rational analysis, involving intensive research on normal and clinical respondents. The dimensions have been shown to be reliable and stable, to have construct validity and to be useful in clinical settings.

For each dimension scores range from 0 to 48, with the distribution of scores in the general population being bell-shaped, and the majority of individuals scoring in the middle of the range, but small percentages at either end. These dimensions can be most readily explained by describing the characteristics of those scoring very high or very low on a particular trait. Descriptions for extreme scores on each of the five dimensions are given in *Table 4*.

It is salutary to realise that although society tends to value a quality such as agreeableness, there are advantages in terms of individual achievements, and indirectly to society, deriving from individuals lacking in agreeableness. Similarly, although extraversion, openness and conscientiousness tend to be viewed positively, if taken to extremes they can in many contexts be a disadvantage. It should be remembered that few respondents will be extreme in their scores, or even if tending towards an extreme score, show all the characteristics listed. It is also important to note that it is normal personality traits that are being measured, so that the neuroticism scale, for example, reflects a dimension of normal personality and is not a measure of psychopathology (such as neurosis). *Table 5* illustrates the distribution of population

TABLE 4 *Meaning of personality traits assessed by NEO personality inventory*

Dimension label	Lowest scorers	Highest scorers
N: Neuroticism	Emotionally **stable**, relaxed even when under stress, even-tempered, secure	Tendency to experience negative feelings, **emotional**, sensitive, maladjusted
E: Extraversion	**Reserved** but not necessarily shy, not exuberant, even-paced, content to be alone	Sociable, **outgoing**, active, assertive, like excitement, talkative
O: Openness	Conservative, prefer the familiar, emotional responses tend to be muted, conventional, '**closed**'	Intellectually curious, imaginative, experience positive and negative emotions keenly, '**open**'
A: Agreeableness	Sceptical of others' intentions, **self-centred**, competitive	Altruistic, **cooperative**, sympathetic to others, eager to help, trusting
C: Conscientiousness	Lackadaisical in working towards goals, less exacting in applying moral principles or behavioural rules, careless	Purposeful, determined, punctual, reliable, scrupulous, overfastidious, workaholic

TABLE 5 *Quartiles for NEO-FFI Form S personality scores for adult women*

Dimension	First quartile	Median	Third quartile
Neuroticism	15	20	26
Extraversion	24	28	33
Openness	23	27	31
Agreeableness	31	34	37
Conscientiousness	32	35	39

scores reported for adult women for the NEO dimensions of personality.[53]

The N (neuroticism) and E (extraversion) scales are potentially most useful in clinical research, in that they may help to explain otherwise anomalous findings. Women with high extraversion scores may be more successful than other women with objectively equal justification for it, in obtaining the healthcare they need or want. This may be because they are better at communicating their medical symptoms and their impact, and so are more likely to be seen by the clinician as warranting the management in question. Alternatively, the appropriateness of the management under consideration may be irrelevant, but they obtain it because they want it and are more assertive than other women. The N score is useful because of the consistent finding that a woman with a high neuroticism personality trait will tend to score higher than less neurotic individuals on any self-report of health complaints, because of a predisposition to be overly sensitive to minor physical sensations and to worry about health. Therefore, the N score can be useful in interpreting self-reported health and experience of healthcare. In addition to these possibilities, it could be predicted that 'open' (O) individuals would want information about their condition and shared decision-making with regard to management, whereas 'closed' individuals would be more accepting of the older ethos that 'doctor knows best'. It might also be anticipated that those with high scores on the agreeableness scale (A) would fit in happily with existing clinic organisation and be trusting of doctors' decisions, whereas those scoring low on this scale would have to be very concerned about their health before risking attendance at the clinic.

Psychiatric health was assessed by an established health measurement scale, Goldberg's General Health Questionnaire GHQ-28[54] (GHQ). The GHQ was originally designed to detect psychiatric disorder in the community, in primary care and among general medical outpatients. In the latter two settings the aim would be to focus on psychological components of ill-health, in case these have relevance to the woman's attendance at the medical clinic. This may either be through having contributed to the decision to consult, or as psychological distress associated with the primary complaint, that needs to be addressed by the clinician. The original use for GHQ involved the application of a cut-off score to identify respondents as 'cases' or not (of psychological ill-health). However, this results in two problems. There is loss of information when converting a potentially broad-ranging score into a binary (case) variable, and the researchers may have no way of knowing the most appropriate cut-off to use (from those that have been published for various applications) for their specific but different study population. One solution to these difficulties is to score the GHQ items on a Likert scale, and then to use for analysis the total of these scores, rather than a case categorisation. This method positions the study population on the entire dimension of possible scores, and hence distributes them relative to each other along the underlying continuum of psychological ill-health. Another potential problem with GHQ is that it has been found that in medical patients there can be some confounding of GHQ psychological ill-health score. This may happen because of the physical symptoms included in the GHQ, or because of the patient's transient anxiety as to the outcome of a clinic appointment to investigate unusual or troubling symptoms. One way to achieve greater insight into the nature of a specific patient's GHQ-measured psychological health is to accumulate the scores for the GHQ items not just into a total score but also separately, for distinct subscales representing subtypes of psychological distress. GHQ-28 has four subscales: A, somatic symptoms; B, anxiety; C, social dysfunction; and D, depression.

Therefore, the scaled version of GHQ, the GHQ-28, was used in this study and Likert scoring (0–1–2–3) was applied. For each subscale the scores can range from 0 to 21, whereas the total score for the entire GHQ can range from 0 to 84. It should be noted that the GHQ focuses on (recent) changes in usual psychological health, rather than chronic conditions. It is well known that for patients referred to many different outpatient clinics there are higher scores on the GHQ.[55] The rationale for measuring psychological well-being in this study was as a potential explanation (psychological co-morbidity) for otherwise anomalous dissatisfaction with the outcome of the clinic appointment. Furthermore, the nature of the co-morbidity can be categorised as: (A) somatic symptoms (which may

indicate an individual sensitivity to physical changes or may be a consequence of the somatic symptoms that often accompany heavy periods); (B) anxiety (perhaps about the cause of the symptoms that have led to the consultation); (C) social dysfunction (which again could result from the effect of heavy bleeding on daily life); or (D) depression.

Women's experiences of endometrial evaluation were assessed prospectively by means of report forms completed immediately after the appointment (Appendix 4). For each randomised investigation undergone, which may have been on that day or later, a separate report was completed immediately afterwards, covering explanation received, time taken and reaction to that investigation. At the end of the initial (recruitment) appointment the woman completed a questionnaire report on her experience of the clinic visit. This included rationale for consultation with the doctor, information received before clinic attendance, prior investigations for abnormal bleeding and time issues. All these questionnaire forms included the possibility of free-text comments.

To ascertain a reflective overall judgement on the woman's clinic attendance and package of investigations, and to allow for the fact that there was sometimes delay to completion of the study ultrasound or hysteroscopy investigations, a further clinic review questionnaire was completed on the day after the last randomised (study) investigation had taken place. This ascertained feelings about the clinic visit and about future health, any experience of abdominal discomfort or bleeding after the clinic appointment, and a judgement as to how worthwhile the clinic visit had been.

Follow-up questionnaires

The women were also asked to complete follow-up questionnaires, sent by mail, at 10 and 24 months. In these they were asked to report whether they still had symptoms, whether, since their initial appointment, they had visited their GP or been a hospital day case or inpatient for the bleeding problem, whether they had attended any hospital gynaecology clinic, how they felt about their care at the time of recruitment, and how they would feel if they required further investigations in the future. The forms used are provided in Appendix 5.

Case-note review

Patients' case notes were reviewed after 2 years (or after 10 months if recruited later on in the study) and details of additional appointments, correspondence, investigations and treatments were recorded. The health economist (BG) was consulted to ensure that the necessary data were collected for the economic evaluation. The case-note review form was devised to standardise the information collected, and to maximise data consistency and reliability. The form used is provided in Appendix 6.

Clinician surveys

To address clinician beliefs and practices with regard to endometrial evaluation, clinicians were surveyed at three points: at the start of the study by interview, and by questionnaire at 23 months into the study, when recruitment was about half way, and again at 40 months into the study, after dissemination to them of the preliminary findings regarding the comparison of performance of the study investigations. Recruitment had ended by the time the clinicians received the report of preliminary findings. A summary of some findings from the surveys is provided in Appendix 7.

Other data collected

Cost data were collected as described in Chapter 3 ('Economic analysis', p. 22) and Appendix 8.

Outcomes

Scope and range of study variables

Given the broad scope of the study there was a wide range of potential outcome variables:

- assigned evaluations: time taken, adverse events, success at measurement/sampling, abnormality/diagnosis
- patient experience: patient rating of experience of evaluations assigned (discomfort, anxiety/distress, reassurance, acceptability), satisfaction at 10 and 24 months (including improvement in presenting symptoms)
- follow-up case-note review, up to 10 and 24 months postrandomisation: number of appointments, adjunctive evaluations and treatments, time to last investigation, time to final treatment decision/discharge
- direct costs to the NHS: calculated up to the stage at which no further investigations took place, up to 10 months postrecruitment and up to 24 months postrecruitment
- clinician attitudes to evaluation: measured among local gynaecologists by survey initially at month 23 of study, and after reading draft report of evaluation findings (month 40 of study), clinician's stated preference for evaluation of each woman recorded on patient details form, pattern of evaluations undertaken across duration of study.

TABLE 6 *Primary outcome variables for study objectives*

Objective	Primary outcome variables	
1	Completion of investigation Adequacy of sample/view Abnormalities detected Occurrence of adverse events After-effects reported by women Clinic visit 'worthwhile'	Compared across the three investigations
2	Adequacy of sample Women's preferences regarding biopsy method	Compared between the two biopsy methods
3	Cost-effectiveness to last investigation	Compared across the three investigations and between the two biopsy methods
4	Medically not possible Completion of investigation Adequacy of sample/view Rated unpleasant by woman	'Failures' modelled on reproductive and demographic factors, and other potential explanatory variables
5	Total resources used and costs	Distributed by time and compared by randomisation option and the woman's characteristics
6	Clinician evaluation preferences	Compared by time, randomisation option and the woman's characteristics

There was also a number of variables that had been measured as potential explanatory variables. These were mainly the woman's characteristics and comprised: stage of cycle at biopsy (calculated from date of LMP), self-report of general health, NEO personality factor scores, GHQ scale scores for psychological well-being, risk group, HRT use, parity, self-stated reason for attendance, previous bleeding complaints, expectation of internal examination, attitude to investigation setting, sensitivity to pain, general tendency to worry about health and concern about bleeding problem specifically.

Primary outcome variables

The main outcome variables are summarised in *Table 6* separately for the different study objectives.

Sample size

Differences between evaluation methods were examined over the study as a whole, as well as within groups. The required sample sizes were calculated to be: high-risk group, 200; moderate-risk group, 400; and low-risk group, 300. With these sample sizes, the important comparison of blind biopsy compared with hysteroscope-aided biopsy would have 360 women in total evaluated by each of the two strategies. As a 5% significance criterion was intended, this sample size ensured 80% power to detect a difference between methods of 7 percentage points (or greater) in the proportion with a specified outcome, for example, adequacy of sample.[56] Similar power was assured for testing the impact of adjunctive ultrasound, where possible TVUS. The within-subject comparison of Pipelle versus brush would have 80% power to detect smaller differences, of 4 percentage points or more. For comparisons within the separate groups power would be lower, but sample sizes were chosen to ensure highest power in the moderate- and low-risk groups, where there was most doubt as to optimum strategy.

Chapter 3
Methods

Randomisation

Sequence generation

Comparison between three methods of endometrial evaluation

Randomisation options within groups were as follows:

- high-risk group: randomisation was to one of two options: hysteroscopy with biopsy, or blind biopsy plus ultrasound (blocked in groups of 40)
- moderate-risk group: every woman in this group received at least biopsy; the remaining two diagnostic procedures (ultrasound or hysteroscopy) were randomly assigned in addition to biopsy, in a factorial design; randomisation was blocked in groups of 80
- low-risk group: these women were randomised to no investigation (40%), receiving standard clinic management by medical treatment and surveillance, or to one of the three investigations (20% to each); randomisation was blocked in groups of 60.

This is summarised in *Figure 1*.

Randomisation of order of the three methods of investigations

Where women received more than one investigation it would be preferable to randomise order, to prevent bias being introduced. However, in the case of hysteroscopy with biopsy the biopsy must be second, or otherwise the hysteroscopic view of the endometrium is likely to be obscured by blood in the cavity, resulting from the preceding biopsy. The alternative would be delays in completing the randomised investigations, and imposition of two clinic visits and invasive procedures.

Similarly, since there was a waiting list for ultrasound, it was accepted that for the women randomised to receive both ultrasound and biopsy (high- and moderate-risk groups) the biopsy should be taken at the initial clinical appointment, when a pelvic examination was being undertaken anyway, and that it would be unethical, and affect clinicians' willingness to allow their patients to be recruited for the study, if any other order was imposed.

Therefore, the only randomisation option that had potential to be randomised for order was

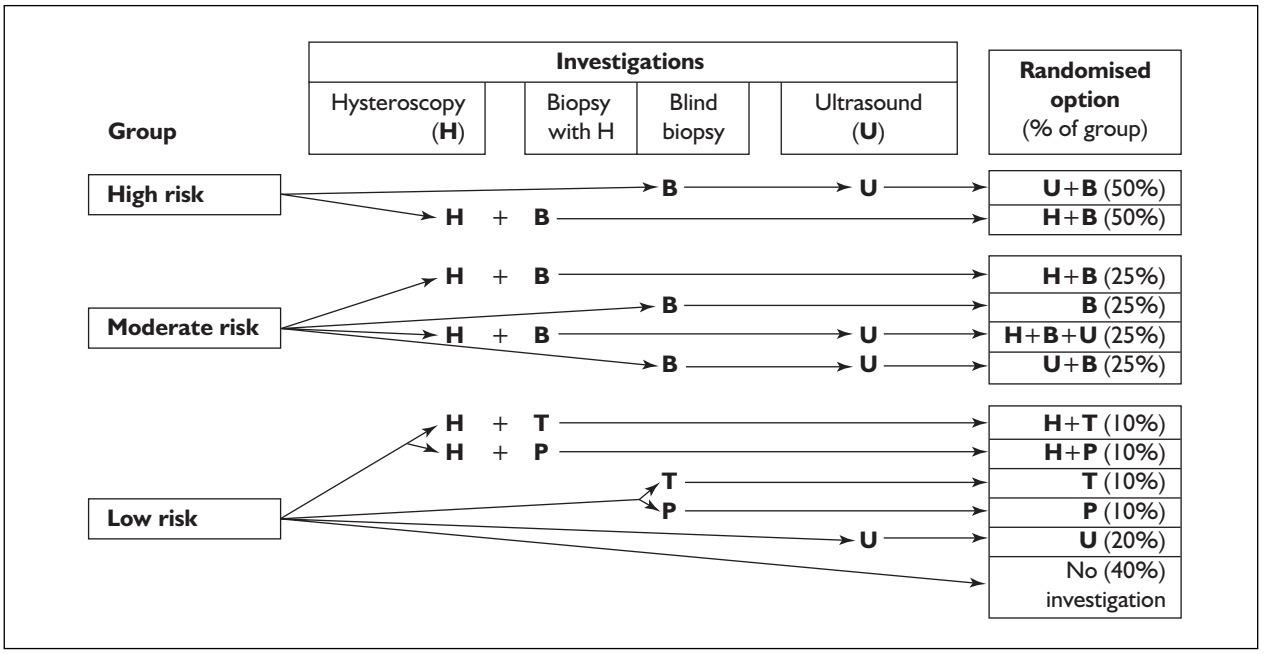


FIGURE 1 *Study design: randomised evaluations by group (ignoring order of investigation). T, Tao sampler; P, Pipelle sampler; B, both biopsies randomised 50% to each order.*

ultrasound plus hysteroscopy with biopsy, in the moderate-risk group. The randomisation specified the order in which these two investigations should be undertaken, with half of the women having one specified first and half the other. However, in the spirit of a pragmatic trial this was interpreted as the preferred order, not an absolute imposition.

Comparison between biopsy methods: Pipelle and Tao brush

For the paired comparison of Pipelle biopsy versus new brush sampler (Tao brush), in the high- and moderate-risk groups, the order in which the two biopsies were to be undertaken was randomly assigned, so that half of the women had one first and half the other. However, in the low-risk group women had only one biopsy, but with random assignment to specific biopsy method: 50% each to Tao brush and Pipelle.

Allocation concealment

Randomisation was undertaken to industry standard via a customised computer program printing onto multipart computer stationery designed specifically for such a purpose. The randomisation codes were printed only on the inside of the sealed envelope, in the manner of payslips, so the precise code could be viewed only if the slip was torn open. The slip was opened only if and when the clinician confirmed that the woman was eligible for the study (and the woman had consented to participation). Shading on the inside of the slip ensured that strong light could not be used to view the contents through the paper.

Implementation

The randomisation envelopes were preprinted, with the master list being kept offsite by PW. Randomisation envelopes were marked with the risk group code letter and numbered sequentially within risk group. They were kept and used in strict numerical sequence within each group. The recruiting research assistants spoke with the women before they were seen by their clinicians. If a woman consented to take part in the study, the next available randomisation envelope for the relevant stratification group (determined by age and menopausal status only) was attached to her recruitment forms. Before the woman was seen by the doctor, the recruiting research assistant described the study to the doctor, gave him or her an information sheet, explained that the woman had agreed to take part in the study, and gave the doctor an eligibility/recruitment form. This was to be completed by the doctor after he or she had spoken to the woman. This form was used to obtain the clinician's consent and, since for premenopausal women under 40 years of age their group could be low or moderate risk, depending on specific clinical risk factors, to confirm the stratification/risk group. If consent was not given by the doctor, or the doctor found that the woman had changed her mind and withdrawn her consent, the allocated randomisation envelope was returned unopened to the stock of envelopes, maintaining numerical sequence.

If the doctor also consented to a patient participating, he or she came out of the room before examining the patient. If the woman had been assigned by recruiting staff to the correct group, the allocated randomisation envelope was opened so that the doctor could see which investigation(s) to carry out, and the order of investigations. (If a Tao brush was needed, the recruiting researcher supplied one to the doctor.) However, if the woman needed to be reassigned to another group (typically from the low-risk group to the moderate-risk group because of clinical risk factors), the envelope tentatively provided was returned unopened to stock as above, and the next available randomisation envelope for the correct group was obtained and opened instead. For the purpose of audit checks the recruiting researcher retained the randomisation assignment once opened, and stuck a hospital patient address label to it, or wrote the woman's name and hospital number on the envelope. The randomisation number became the woman's study number and was written on all of the woman's study forms, including the questionnaires still to be completed by the woman, on the log sheet to note recruitment, and on the assignment envelope which was kept with study administration records and consent forms.

The recruiting researcher waited while the investigations were carried out, and labelled any study samples appropriately before they were sent off for analysis. If the woman was randomised to hysteroscopy or ultrasound, the recruiting researcher arranged this with the woman immediately after her clinic appointment. If the woman was unable to attend a specific ultrasound session, the recruiting researcher put a study sticker on the woman's ultrasound request card.

Blinding

The nature of the interventions (their being procedures undertaken by the clinician and

undergone by the woman) meant that blinding was not possible.

Statistical methods

Enhancing quality of measurement

Reliability check on case-note review

The case-note review was a complex task, requiring careful scrutiny of notes and judgements as to relevant data to be transcribed. An initial policy of overinclusiveness was adopted, and after a number of case-notes had been reviewed the process was discussed. After discussion of the data extracted by the relevant research team (the two case-note review staff, a senior clinical investigator and two statisticians) a protocol was devised to focus data collection on important and relevant aspects of the case notes, and to guide selection and transcribing. This process of data extraction, discussion of problematic cases and refinement of protocol, was repeated two or three times.

Next, a reliability check was undertaken, with a set of 12 notes (four from each risk group) being selected at random from those already reviewed by one or other reviewer, and then subjected to independent and blind data extraction by the other case-note reviewer, and by the two statisticians, following the agreed protocols. The aim was to check inter-rater reliability between the two reviewers, and by involving the two statisticians, who had been party to all the discussions but who had not heretofore reviewed notes, to ascertain the extent to which the written protocol enabled data to be extracted from the notes in a manner consistent with the two practised and trained reviewers.

This exercise ensured that the data extraction protocol could be applied satisfactorily by research workers new to the case-note review, but even more importantly, given the small number of minor differences found, ensured that the two study reviewers were consistent.

For the remainder of the case-note review the policy of overinclusiveness of data for difficult or anomalous cases was continued, to minimise the risk of bias. These cases would then be discussed by the research team. Wherever possible this discussion was undertaken blind to the randomised investigation(s) or the outcome, or both, as long as the point at issue was not one or other of these aspects. Wherever a new decision was made the case-note review protocol was updated to include it for the future.

Data coding and checks

All forms were checked by the study data manager for errors or omissions before being passed to a skilled data entry clerk who entered the data into a purpose-designed access database. The database was designed to undertake logic and range checks as the data were entered, and the clerk (who was part of the research team and so developed a good understanding of the study aims and methods) also marked any anomalies identified. These were checked, action was discussed by the research team if necessary, and they were corrected. Finally, once data had been entered the entire data set was examined and checked by the statistician, and any errors or anomalies (e.g. with dates) were rectified if the information necessary could be recovered.

The case-note review involved extraction of a total of 10,722 events (e.g. letters, appointments, investigations, treatments) from the case notes. These data constituted a huge coding task for the study data manager, with many of the coding decisions having to be discussed by the research team, including the senior clinical and statistical investigators (HODC and PW), the statistician (AJL) and often the health economist (BG). The coded case-note data were then entered into the database with corresponding dates, so numerous date checks had to be programmed into the data entry database. A complex and hence extensive SAS 8.02 program then had to be designed to match codes to established costs and accumulate the various total costs for each woman (as determined by the various cut-offs specified by the analysis plan).

Recoding of women's responses on reports and questionnaires

Investigation reports

To simplify the reporting of the study results it was necessary to recode some of the variables. This was particularly so for the patient acceptability and comfort items. Each woman was asked to rate the investigation she had undergone on the following scale: fine, slightly unpleasant, quite unpleasant, very unpleasant, and extremely unpleasant. Responses of 'very unpleasant' or 'extremely unpleasant' were combined and presented labelled as markedly unpleasant. Each woman who had a biopsy was asked whether they suffered from any abdominal pain ('cramps') or discomfort with the procedure; the responses were: not at all, minimal, a little, a lot, and severe. Responses of 'a lot' or 'severe' were combined and presented labelled as having suffered markedly with abdominal pain or discomfort.

Review of clinic attendance

Women were asked to record their feelings about their clinic attendance by rating six statements as 'not true at all', 'not very true', 'fairly true' or 'very true'. Agreement to a statement was defined as a 'fairly true' or 'very true' response. Women were also asked how they felt about their future health with respect to their bleeding problem.

Assumptions regarding missing data

For high- and moderate-risk women who were all randomised to both types of biopsy the order of undertaking the biopsies was not recorded for three women and it was different to the randomised order for a further 18 women. For four of these latter cases, further information was available and it was clear that the biopsies had been carried out in the reverse order. For the remaining 17 biopsies, the order was set to missing as it could not be confirmed whether the clinician ticked the wrong box on the form or whether the biopsies were indeed completed in the reverse order. In these cases the women's responses regarding the preferred biopsy were changed to 'missing' unless the woman expressed 'no difference between the two biopsies' (since the question was asked in terms of a preference for the first or second biopsy). Of the four women where the biopsies were completed in the reverse order, the preference was changed for three women to reflect the actual order rather than the randomisation order. For the remaining woman the biopsies were attempted in randomisation order, but the first biopsy was initially unsuccessful and was reattempted after the second biopsy. However, this woman expressed no preference between the first and second biopsies so the response was unchanged for the analysis.

Parity was noted only on the biopsy form or hysteroscopy (with biopsy) form. Therefore, parity was not ascertained for women randomised to ultrasound alone, or to no investigations, for 32 and 62 low-risk women, respectively. Furthermore, in some cases, even when there was randomisation to biopsy, parity was not entered on the biopsy form, perhaps because the biopsy failed. Of 297 women who were randomised to biopsy at hysteroscopy, parity was not noted for 17 women (six high-risk, eight moderate-risk and three low-risk women), and of the 292 women who were randomised to blind biopsy, parity was not noted for 24 women (13 high-risk, nine moderate-risk and two low-risk women). Given that within the low-risk group there was randomisation to investigations, one would anticipate that the estimates for nulliparity derived from women having biopsy should also apply to those randomised to the ultrasound and no investigation options.

Definition of failure for investigation

Results for the three investigations were categorised in different ways to compare the success or failure of the investigations. The first failure categorisation related to whether the investigation was completed. The responses on the relevant forms were coded as shown in *Table 7*, to indicate completion or failure.

The second failure categorisation related to whether or not the investigation was judged to be a success by the operator (hysteroscopy or ultrasound) or the pathologist (biopsy). The outcomes that classified the investigations as successful or not are shown in the *Table 8*. Although diagnostic ability would be expected

TABLE 7 *Coding of investigations as completed or not*

Investigation	Completed	Not completed
Biopsy	Done	DNA/medically inadvisable/abandoned[a]
Ultrasound	Done	DNA/medically inadvisable
Hysteroscopy	Done	DNA/medically inadvisable/abandoned[a]

[a] Often for failed insertion.
DNA, did not attend.

TABLE 8 *Coding of completed investigations as successful or not*

Investigation	Successful (view/sample)	Unsuccessful
Biopsy sample	Adequate/barely adequate	Inadequate/no sample
Ultrasound	Good view/difficult view	No view
Hysteroscopy	View obtained	View obscured/too dark

to be less for 'barely adequate' samples and ultrasound views that were 'difficult', the pathologist (AW) and radiographer (SC), respectively, considered that these should nevertheless be categorised as 'successes'.

When comparing the sample quality between the Pipelle and Tao brush biopsy methods, all women for whom at least one sample was obtained for either method, were used in the analysis. For example, if failed insertion occurred for the Pipelle and a sample was obtained for the Tao brush, then the Pipelle sample was classified as a failure.

Statistical tests

Wherever feasible the comparisons of the main study outcomes will be on the basis of ITT, using as the denominator all women assigned to that randomisation option (N_t). This is considered best practice in analysis of randomised trials, to avoid inadvertent bias being introduced by factors associated with completion of the assigned procedures.[56] However, for clinical insight descriptive results are also on occasion presented using as the denominator other relevant n, such as the number for whom the test is medically potentially possible (N_m), the number for whom it has been possible to complete the test (N_c), and the effective n, defined as all cases with complete (recorded) data for variables used in a specific analysis (N_e). This allows a clinician to ascertain, say, the rate of adverse effects that may be expected among women in whom the test is actually undertaken.

Where there are missing binary data for some women in the subgroup, then an outcome for them has to be imputed as a positive outcome or a negative outcome. This will be explained in the accompanying text. In elaborative analyses, for example looking at the factors associated with marginal improvements in outcome, the effective n will be used.

The chi-squared (χ^2) or Fisher's exact test was to compare proportions among groups. Spearman's non-parametric correlation coefficient rho was calculated to assess association in ordinal variables. McNemar's χ^2 test (with the continuity correction) was used to compare the proportion of successful biopsies between the two types of biopsy (paired data). McNemar's odds ratio (OR_{McN}) was also calculated with 95% CI. The OR_{McN} gives the likelihood that among discordant pairs an inadequate sample comes from the Pipelle, so an OR_{McN} of 10 for a particular group of women would imply that among discordant results in that group the number of inadequate samples for Pipelle was ten times the number for the Tao brush.

Stepwise binary logistic regression was used to predict which women were more likely to suffer from cramp, bleeding and discomfort. For these analyses the after-effects variables were converted to binary form by combining women's responses of 'some', 'a lot' or 'severe' (as experiencing the after-effect) and 'none at all' or 'hardly any' (as not experiencing it). These analyses are reported in terms of OR and CI for the OR. An OR of 1 means that there is no difference in odds of after-effects between subgroups based on the predictor variable in question, an OR greater than 1 implies that there is an increased odds (the variable is a risk factor for after-effects) and an OR less than 1 implies that there is decreased odds (the factor is protective against after effects). For example, if the OR is 1.2 for age then the odds of after-effects increases by a multiple of 1.2 (20% increase in odds) for each additional year in age. The stepwise procedure adds (and removes) variables to the model depending on their statistical significance, so in the final model the variables are jointly all significant independent predictors of increased or decreased risk of after-effects. The variables considered for entry into the logistic regression were age, whether the woman was randomised to hysteroscopy, ultrasound and/or biopsy/biopsies, risk group, GHQ subscale totals (A, B, C and D), NEO inventory scale totals (for factors N, E, O, A and C), current health and whether the woman worries about her health. The Kruskal–Wallis test was used to compare the time taken to perform the investigation and to compare costs and the NEO and GHQ scale scores among risk groups.

Boxplots have been produced to illustrate the distributions of the NEO and GHQ components among the three risk groups, and have also been used to illustrate the distributions of skewed data such as costs. Boxplots show the median (solid line), upper and lower quartiles (bottom and top of box) and the general spread of the observations (whiskers and outlying values). Any observations that are between 1.5 and 3 times the box length away from the end of the box, on that half of the distribution, are classified as outliers, and any more than 3 box lengths away, as extreme values. Outliers are shown as open circles and extreme values as stars. The whisker at each end of the box

joins the box end to the farthest value (if there is one more extreme than the relevant quartile) that is neither an outlier nor an extreme value.

Economic analysis

Question

An economic evaluation was conducted during 2000–2001 alongside a randomised trial comparing a range of methods of endometrial evaluation, to address the question: What is the most cost-effective method for outpatient evaluation of abnormal uterine bleeding?

Selection of alternatives

The investigation methods being compared – blind biopsy, hysteroscopy (with biopsy) and transvaginal/abdominal ultrasound – are currently in routine use in the NHS. Women attending gynaecology outpatient clinics with AUB were placed into groups on the basis of age/menopausal status and risk factors for endometrial cancer: high risk, moderate risk and low risk. The evaluation options to be compared differed by group, and in the two higher risk groups involved combinations of investigations, to ensure that all women had adequate investigation of their symptoms, considering risk status. In the low-risk group one of the randomised options was 'no investigation'. In the high-risk and moderate-risk groups two methods of biopsy were compared in every woman: Pipelle and a newer device, the Tao brush.

Form of evaluation

The economic evaluation took the form of a cost-effectiveness analysis and the perspective adopted was that of NHS Scotland. Cost-effectiveness of the various methods was compared, with the outcome measure being no further investigation being required. Cost per patient was calculated for each comparator. All hospital consultations (outpatient department, day case and inpatient), treatments (medical or surgical, e.g. endometrial ablation), investigations, tests, letters and prescribed drugs were costed. Further data on treatments following the establishment of a working diagnosis were collected to reflect cost consequences for investigation that allowed discovery of additional pathology (e.g. fibroids or polyps) compared with others.

Effectiveness data

Information from the case-note review at 2 years from the start of recruitment (January 1999) was used to identify the point at which no further investigations were carried out, on the premise that investigations were carried out until a satisfactory diagnosis was obtained. In this way all methods of evaluation were costed until they yielded the same outcome.

Costing

Care was taken in this study to collect detailed information to allow cost estimation for all relevant items of service. The economic evaluation was carried out according to recognised guidelines.[57] In particular, the costing methodology was in accordance with Graves and colleagues,[58] who published critical comment on the quality of cost methods used to derive patient-level costs in economic evaluations conducted alongside randomised controlled trials.

Unit costs were estimated in UK pounds in financial year 2000/01, with all cost data taken from earlier years being uprated according to the British Medical Association cost/price index.[59] Unit costs for medically qualified staff include an element of ongoing training, which would cover the introduction of new technologies.[60] Total costs were calculated for each patient following initial endometrial evaluation, and included costs of the procedure itself, outpatient attendances, additional investigations (outpatient and inpatient), drugs and other treatments.

Costs for investigations and tests were specially calculated if specific information on them was known. The information was used to calculate costs which were then added to the following base costs for gynaecology:[61]

- consultant outpatient appointment of 10–15 minutes: £57
- day case: £294 (excludes £116 for theatre and £26 for laboratory tests, but includes £46 for pharmacy)
- inpatient: £278 per day (excludes £251 per stay for theatre and £97 per stay for laboratory tests, but includes £25.30 for pharmacy).

Where information on investigation and tests was not known, specific costs were not calculated, and instead average gynaecology speciality costs were used:

- day case: £436 (includes theatre, laboratory tests and pharmacy)
- inpatient: £1182 (includes theatre, laboratory tests and pharmacy) for a length of stay of 3 days (or not known).

TABLE 9 *Costs used or calculated for study randomised outpatient investigation options*

Randomised outpatient investigations	Cost[a]
(Blind) biopsy by Pipelle sampler only[b]	£64.68
(Blind) biopsy by TAO brush only[b]	£93.99
(Blind) biopsy by both Pipelle and TAO[b]	£158.68
Hysteroscopy[d] + Biopsy by Pipelle only[c]	£190.53–248.14
Hysteroscopy[d] + Biopsy by TAO only[c]	£219.84–277.45
Hysteroscopy[d] + Pipelle and TAO biopsies[c]	£273.44–331.05
Transvaginal/abdominal ultrasound[c]	£42.53

[a] These are the additional costs for the randomised investigations assigned by the study. To obtain the total cost for a specific combination of randomised investigations, the base cost for the initial gynaecology clinic appointment (£57) needs to be added to the cost for ultrasound, if assigned, and the relevant cost for the assigned biopsy combination (with or without hysteroscopy). For women randomised to no investigation there are no additional costs due to the study, so the total cost is £57.

[b] These are the costs solely for the relevant study biopsy (or combination of biopsies), and do not take into account the base cost for the clinic appointment required to be able to undertake them (£57).

[c] These costs also are in addition to the initial gynaecology clinic appointment. However, they include further outpatient clinic costs, as relevant, because hysteroscopy and ultrasound are separate outpatient events.

[d] Hysteroscopy costs are shown as a range where the lower cost represents the cost of using a hysteroscope with a reusable (autoclavable) sheath and the higher cost represents the cost of using a hysteroscope with a disposable sheath.

Table 9 shows the costs calculated and used for study investigations. All other costs calculated and used are presented in Appendix 8.

Modelling adjustments for timing of costs and benefits

All unit costs were adjusted according to published healthcare cost/price indices to 2000/01 values if from years prior to that.[60] When costing capital equipment, annuitisation methods were applied to calculate an equivalent annual cost (utilising a discount rate of 6% per annum).

Allowance for uncertainty

A number of assumptions had to be made during the course of the calculations for the economic evaluation. A method devised to test the sensitivity of the results of the economic evaluation to variations in values that are surrounded by a certain amount of uncertainty is sensitivity analysis.[51]

A key area explored, where uncertainty exists, included the costing of the use of a reusable, autoclavable hysteroscopy sheath, as opposed to sterile, disposable hysteroscopy sheaths. This means that it will be possible to calculate comparative costs for disposable sheath hysteroscope use. However, reusable sheaths were used throughout the trial; therefore, any savings that are possible owing to reduced infection rates will not have been captured in the trial; however, the additional unit cost of using the disposable sheaths has been calculated.

The outcomes of the study were well recorded in the patient notes, thereby creating less uncertainty than the costs. To allow for a range of uncertainties, total costs per person were recalculated with the cost of each type of evaluation method increased or decreased by 10%; the costs of individual laboratory and diagnostic tests and drug therapies used were of less impact to the total costs per person, so were not varied under sensitivity analysis.

As the comparisons need to reflect cost to the health service, where one of the randomised options was not undertaken (for medical reasons, or because the woman withdrew from the study or failed to attend) the cost of that investigation was not added to the woman's study investigation costs.

There is a problem in comparing the cost-effectiveness of Pipelle and Tao in the high- and moderate-risk groups. This is because of the paired design where every woman has biopsy by both methods. In such circumstances, if one device tends to perform poorly and the other well, the deficiencies in the performance of the first have no observable cost consequences, since the failure to obtain an adequate sample by the less effective device is compensated by the sample and histological analysis from the other device. Therefore, the clinician does not have to decide on subsequent strategies to make good the failure of the first device, as would need to happen in usual clinical practice where only one biopsy is

undertaken. To try to make some estimate of the cost consequences for a failure in one device, if it had been the sole method applied, within the high- and moderate-risk groups the cost difference was calculated between women where an adequate histological analysis was achieved, one way or another, and those for whom neither device produced an adequate sample. (The latter were mainly women for whom insertion failed.) This difference in cost could be used as an estimate of the cost consequences of failed sample where only one biopsy device is used. It will, however, be an overestimate for women for whom insertion is possible, where the failure would have been simply one of adequacy of sample, since for them one of the options available in clinical practice would be then to undertake biopsy using the other, more efficient device. This is discussed further in Chapters 4 and 5.

Presentation of results

The design allowed the most cost-effective method within each risk group to be identified. Results were examined in the base case and under the sensitivity analysis scenarios (where the cost of each evaluation method was ±10% in turn, while holding all other values constant).

Reporting of diagnostic performance

The Standards for Reporting of Diagnostic Accuracy (STARD) initiative aims to improve reporting of studies of diagnostic accuracy, and recommends the use of flow charts and check lists.[62] This was considered for the present study, but not adopted for a number of reasons. The present study was not primarily a comparison of accuracy of diagnosis of endometrial cancer, but aimed to compare cost-effectiveness and acceptability. It was not powered to include enough cancer cases to enable rigorous comparison of sensitivity and specificity across investigation methods. As it was a pragmatic trial it was not designed to ensure gold-standard reference tests for every woman evaluated for AUB, that is, another test that detects and excludes cancers perfectly. However, the study adhered to the principles of reporting espoused by STARD, with full details on participants and recruitment, completion of tests and reasons for non-completion, description of reference test(s) used and estimates of diagnostic accuracy with confidence intervals.[62]

In the case of endometrial cancer, there is no non-invasive test that can be considered to be gold standard. However, each individual study investigation in a particular woman can be compared with all other randomised investigations completed for that woman, and the results of any other subsequent examinations where cancer could be detected (e.g. endometrial curettage). These tests jointly can be the 'reference' test for endometrial cancer. However, in certain cases, a reference test may not have been completed (e.g. a failed insertion for a biopsy but insufficient clinical concern to proceed to endometrial curettage).

When a test has been completed, the result may be abnormal, normal or inconclusive. Inconclusive index test results may miss cancers (detected by other successfully completed investigations), so they have been included (as 'negative' test results) in estimations of diagnostic accuracy. This is to ensure that sensitivity is not overestimated. This strategy may slightly overestimate specificity, so specificity estimates have been provided both including and excluding women with no cancer detected by reference test(s) and inconclusive index test results.

Chapter 4
Results

Recruitment and randomisation

Duration of recruitment and follow-up

Recruitment into the study started in January 1999 and continued until May 2001. The clinic visit review questionnaire was given to each woman at recruitment with a prepaid envelope and women were asked to complete and return the questionnaire the day after their last study investigation. The 10-month and 2-year follow-up postal questionnaires were sent to women 10 months and 2 years, respectively, after recruitment into the study, with one reminder sent a month later if the questionnaire had not been returned. However, for the 2-year follow-up women recruited in August 1999 or later were not sent questionnaires, since their 2-year anniversaries fell outside the data collection phase of the study. Rather than draw case notes twice for review (at 10 months and at 2 years), a single combined review was undertaken at least 2 years following recruitment, extracting data for both follow-ups. However, for those women for whom the 2-year anniversary would be too late (as explained above), their case notes were reviewed before the end of data collection, but only for the shorter 10 month follow-up.

Participation rate and randomisation

A total of 1767 women was assessed for eligibility for the study, and if a woman was deemed eligible the study was described to her and she was invited to participate. There was a high recruitment rate, with both the woman and clinician agreeing to the woman's participation in 67% of eligible women.

As detailed in the methods, women were divided into three risk groups based on their age, menopausal status and risk factors. The groups and their recruitment targets were:

- high-risk group (postmenopausal): 200 women
- moderate-risk group (premenopausal aged ≥ 40 years, or aged < 40 years with risk factors): 400 women
- low-risk group (premenopausal aged < 40 years and without risk factors): 300 women.

Table 10 summarises the number of women assessed for eligibility, deemed eligible and recruited. Detailed participant flow diagrams for the three risk groups are given in the figures in Appendix 9 in the recommended format for Consolidated Standards of Reporting Trials (CONSORT).[63]

The recruitment target of 200 for the high-risk group was reached, but fewer women participated in the study in the younger age groups: in the moderate- and low-risk groups 326 (81.5% of the target) and 157 (52.3% of the target), respectively. Despite the lower numbers recruited in the lower risk groups, the proportion of women assigned to each investigation option was similar to the proportion planned, as blocked randomisation was used (*Table 11*).

Every effort was made to ensure that if there were two biopsies they were undertaken in the order as randomised. However, given clinic organisation and waiting lists for appointments it was neither ethically nor pragmatically feasible to enforce the randomised order for ultrasound versus other assigned investigation, for the three subgroups where this was theoretically possible. Therefore, the randomisation order for investigations other than biopsy has been ignored for analysis and reporting. The clinic organisation meant that in practice blind biopsy was usually undertaken at the initial clinic appointment, and therefore preceded ultrasound, as it did in 97% of women randomised to ultrasound plus biopsy. The ordering of

TABLE 10 *Eligibility and agreement to participate*

Risk group	Assessed for eligibility	Eligible for participation	Recruited	Participation rate (%)
High	459	298	200	67.1
Moderate	761	469	326	69.5
Low	547	260	157	60.4
Overall	1767	1027	683	66.5

investigations in women randomised to hysteroscopy (with biopsies) plus ultrasound reflected the availability of appointments for the two visualisation procedures, with ultrasound second in 56% of cases overall, but with a shift across the time-course of the study so that for approximate thirds of the study ultrasound came after hysteroscopy in 66% of the first third of cases, 57% of the middle third and 39% of the final third.

Recruitment, completion of assigned investigations and study measures, and participant flow through the study

Investigations completed

In practice it was not possible to complete every randomised investigation. *Figure 2* presents the numbers of women recruited into each

TABLE 11 *Randomisation*

Risk group	Investigation	Required *n* (% of group)	Actual *n* (% of group)	Biopsy order: P first:second	Investigation order: U first:last
High	PT + U	100 (50%)	100 (50%)	50:50	50:50
	(H + PT)	100 (50%)	100 (50%)	50:50	NA
Moderate	PT	100 (25%)	80 (25%)	40:40	NA
	PT + U	100 (25%)	80 (25%)	39:41	40:40
	(H + PT)	100 (25%)	84 (26%)	43:41	NA
	(H + PT) + U	100 (25%)	82 (25%)	40:42	40:42
Low	No evaluation	120 (40%)	62 (39%)		
	P	30 (10%)	17 (11%)		
	T	30 (10%)	15 (10%)	NA	NA
	(H + P)	30 (10%)	17 (11%)		
	(H + T)	30 (10%)	14 (10%)		
	U	60 (20%)	32 (20%)		

H, Hysteroscopy; U, transvaginal ultrasound; P, Pipelle; T, Tao brush; NA, not applicable (either one biopsy only or no biopsy, or ultrasound alone or no ultrasound).

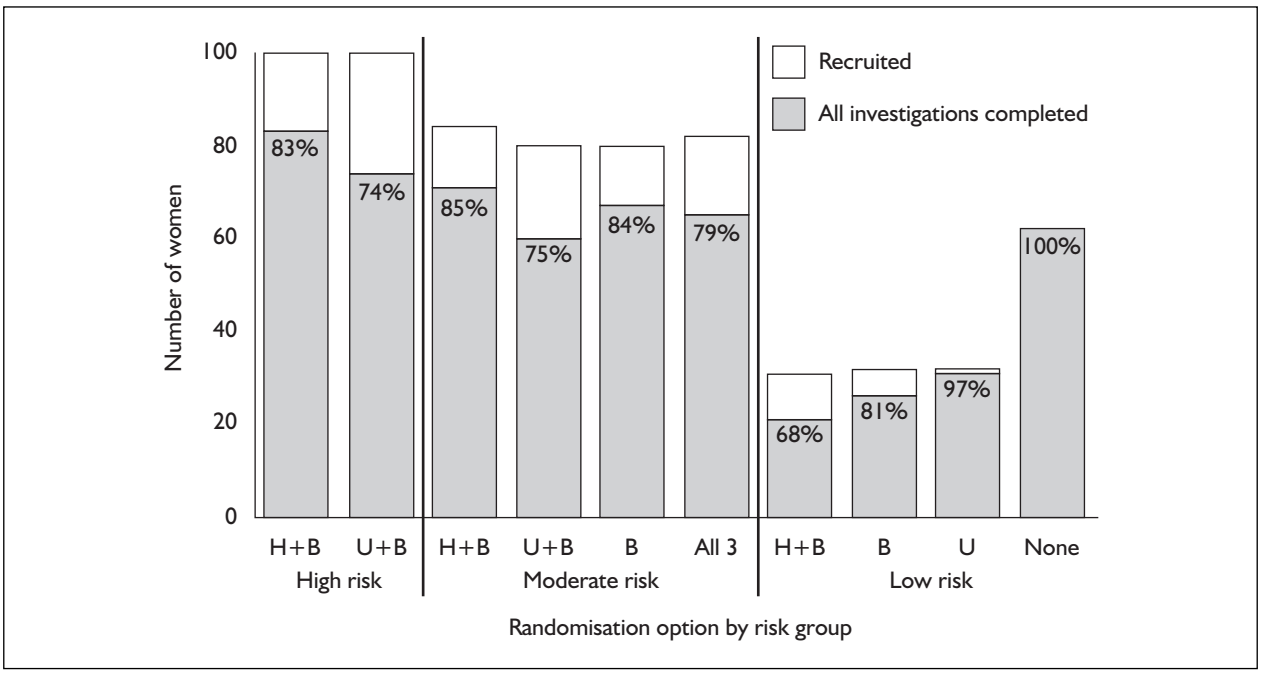


FIGURE 2 *Histogram by randomisation option within risk group, of number of women recruited and the subsets with all study investigations completed (annotated with percentage completed). H, hysteroscopy; U, transvaginal ultrasound; B, biopsy.*

randomisation subgroup, and the numbers having all investigations completed successfully.

Table 12 gives the total number of women who had their randomised investigations completed separately for each investigation and by randomisation subgroup. The overall proportion of successfully completed investigations was 82%. For each woman who did not have all their investigations, the flowcharts (Appendix 9) give the reasons why this was not possible.

In a number of cases where ultrasound was 'completed', this was abdominal ultrasound only, because a good enough view had been obtained with abdominal ultrasound and the transvaginal approach was not required, or because it was not possible or there was insufficient time. Completion figures are presented for all ultrasound investigations, and for those who had TVUS specifically. In some cases the type of ultrasound applied was not recorded (for 6%, 11% and 3% of ultrasound investigations completed in the high-, moderate- and low-risk groups, respectively). The percentages of completed ultrasounds where it was recorded that TVUS was used were 82%, 59% and 68%, respectively, in the three groups. Most often women had both abdominal ultrasound and TVUS (76%, 57% and 65% of completed ultrasounds in the three groups, respectively).

Questionnaires, reports and forms completed at the time of recruitment/investigation

Clinician-completed study forms

Table 13 shows the numbers of women in each arm for whom clinician-completed forms were received regarding the woman's details and, where applicable, performance of assigned investigations and pathology results.

Patient-completed questionnaires and reports

At the recruitment appointment women were given the Health questionnaire, the NEO personality inventory and the GHQ to complete. It was intended that these questionnaires be completed before the consultation with the clinician, but clinic organisation and pressures of time meant that this was not always possible. Since some of the items addressed anticipation of investigations and worries about symptoms, there is the possibility that completion after consultation could affect responses. The Health questionnaire therefore asked the respondent whether the questionnaire had been completed at the clinic or at home, so that checks could be made for different patterns of response. In the event, 32% of Health questionnaires were completed afterwards, at home, and the remaining 68% at the clinic, which in the vast majority of cases was before consultation with clinician.

TABLE 12 *Investigations successfully completed*

Risk group: randomisation	Number (%) of women							
	Total	Hysteroscopy (H) done	Tao (T) biopsy done	Pipelle biopsy (P) done	Both biopsies (B) done	Ultrasound (U) done		All done
						Any U	TVUS	
High								
H B	100	84	89	89	89	–	–	83 (83%)
UB	100	–	75	75	75	95	78	74 (74%)
Moderate								
HUB	82	72	70	71	70	75	46	65 (79%)
H B	84	71	75	76	75	–	–	71 (85%)
UB	80	–	64	64	63	77	43	60 (75%)
B	80	–	69	69	67	–	–	67 (84%)
Low								
H T	14	11	12	–	–	–	–	10 (71%)
H P	17	13	–	12	–	–	–	11 (65%)
U	32	–	–	–	–	31	21	31 (97%)
T	15	–	12	–	–	–	–	12 (80%)
P	17	–	–	14	–	–	–	14 (82%)
None	62	–	–	–	–	–	–	62 (100%)

TABLE 13 *Numbers of forms completed by clinicians and pathologists*

Risk group	Randomised investigation	*n*	Clinician-completed forms				Pathologist-completed forms[b]	
			Patient form	Investigation forms[a]			Pipelle biopsy	Tao brush biopsy
				H (+B)	B	U		
High	PT + U	100	100	–	99	95	73	72
	(H + PT)	100	100	94	–	–	86	89
Moderate	PT	80	80	–	79	–	70	68
	PT + U	80	80	–	80	77	64	64
	(H + PT)	84	84	82	–	–	75	76
	(H + PT) + U	82	82	77	–	75	71	69
Low	No evaluation	62	62	–	–	–	–	–
	P/T	32	32	–	30	–	14	12
	(H + P/T)	31	31	28	–	–	12	12
	U	32	32	–	–	31	–	–

[a] In most cases the clinician completed the report even when an investigation failed.
[b] The maximum potential *n* for pathologist forms is shown under the B or H+B column. The actual number of forms due is less the number of biopsies, as some biopsies obtained no sample. In the high- and moderate-risk groups each of Pipelle and Tao have the indicated potential *n*, whereas in the low-risk group only one or other biopsy was done, so Pipelle and Tao *n*s total this potential *n*.
–, no report required.

TABLE 14 *Numbers received of patient-completed recruitment questionnaires, and reports on clinic visit and assigned investigations*

Risk group	Randomised investigation	*n*	Recruitment Qs			Reports on clinic			
			Health Q	NEO	GHQ	Clinic report	Investigation reports		
							H	B	U
High	PT + U	100	95	91	94	95	–	85	90
	(H + PT)	100	91	87	90	89	84	82	–
Moderate	PT	80	73	72	74	74	–	73	–
	PT + U	80	79	77	79	79	–	72	74
	(H + PT)	84	82	80	82	82	79	82	–
	(H + PT) + U	82	80	77	80	80	72	76	73
Low	No evaluation	62	56	54	56	56	–	–	–
	P/T	32	29	28	29	29	–	26	–
	(H + P/T)	31	26	25	25	26	26	26	–
	U	32	29	28	29	29	–	–	27

In some cases a woman who had a failed investigation nevertheless completed the report on it.
Q, questionnaire.

Table 14 shows the numbers of women in each arm completing the relevant recruitment questionnaires, and there-and-then reports about the clinic visit and randomised investigations. Numbers completing the later clinic review and follow-up questionnaires are detailed separately for risk groups in the participant flow diagrams (Appendix 9, *Figures 23–25*).

Participant flow

The figures in Appendix 9 present separate flowcharts for each of the three risk groups, detailing participants' progress through the phases of the randomised trial. The charts give the number of women assessed for eligibility for the trial, the reasons women were excluded before randomisation, and details of those who did not

TABLE 15 *Characteristics of participating women, by risk group*

Characteristic	Risk group		
	High (*n* = 200)	**Moderate (*n* = 326)**	**Low (*n* = 157)**
Mean (SE) age (years)	57.6 (0.57)	45.2 (0.26)	33.9 (0.34)
Age (years) (%)			
19–29		1	11
30–34		2	37
35–39		3	52
40–44	1	36	
45–49	9	40	
50–54	35	17	
55–59	23	1	
60–64	15		
65–69	7		
70–79	10		
80–86	1		
On oral contraception (%)	1	3	15
Sterilised (%)	22	38	28
On HRT (%)	30	9	0
Presenting complaint[a] (%)			
Postmenopausal bleeding	95	1	0
Heavy periods	1	68	57
Postcoital bleeding	2	8	10
Intermenstrual bleeding	2	22	27
Irregular periods	5	47	46
Other bleeding problems:			
Bleeding on tamoxifen	1	0	0
Pain	1	3	5
Long periods	0	2	3
Frequent periods	0	1	1
Other	0	2	1

[a] Note that women could answer 'yes' to more than one complaint, so the sum will total more than 100%.

have their randomised investigations or were ineligible for further follow-up, or both.

Baseline data

Participants' characteristics

Totals of 200, 326 and 157 women, respectively, were recruited into the high-, moderate- and low-risk groups. Owing to the nature of the different risk groups, some of the women's characteristics were expected to differ between the three groups, as illustrated in *Table 15*.

For virtually all women categorised as moderate risk (premenopausal and aged ≥ 40 years, or < 40 years but with risk factors) it was because they were at least 40 years of age (308, 94%), rather than because of having risk factors (polycystic ovarian syndrome, prior use of unopposed oestrogens/ tamoxifen, obesity, diabetes and family history of endometrial cancer). Of the remaining 18, nine women were obese and seven had polycystic ovarian syndrome (and two of these were also obese). However, for the remaining four women the nature of the risk factors determining the status as moderate risk was not recorded. A small number of erroneous group assignments was made, with six women being switched between low and moderate groups, three each way. A further two women were wrongly assigned as high risk, when one was low and the other moderate risk. Since randomisation options were very similar across the three risk groups, and there was no bar to undertaking additional investigations as wished, it was judged that these were genuine errors and not deliberate violations to protocol with adverse impact on validity.

Additional clinician-reported characteristics are given in *Table 16*. The highest rate of nulliparity was in the youngest group (31%). The distribution by day of menstrual cycle was as expected for the

TABLE 16 *Clinician-reported patient characteristics of potential relevance to outcomes, separately by risk group*

Characteristic	p^a	Risk group		
		High (n = 200) (%)	Moderate (n = 326) (%)	Low (n = 157) (%)
Nulliparous (n = 181, 309, 58)	–	13	15	31
Previous bleeding complaints				
Heavy periods	0.001	5	23	13
Postmenopausal bleeding	0.001	8	1	0
Irregular periods	0.001	4	12	14
Phase of cycle at biopsy[b]				
days 0–7		7	16	17
days 8–16	0.001	3	25	24
days 17–22		13	14	15
days 23–40		16	21	30
> 40 days		61	25	15
Clinician investigation preference(s)				
No evaluation	0.001	0	4	24
Biopsy	0.001	92	85	25
Hysteroscopy	0.004	17	25	13
Ultrasound	0.001	75	53	48
Other	0.001	5	5	20

[a] χ^2 test for comparison of proportions.
[b] Calculated from date LMP – on account of missing data n = 31, 288, 54 respectively.

31 postmenopausal women, providing LMP is taken to mean withdrawal bleed on HRT. The high proportion of women in the two lower risk groups who at biopsy were more than 40 days past the start of their LMP (as recorded at recruitment) may be 'biopsy with hysteroscopy' women who have had to wait for the hysteroscopy and in the meantime have had an intervening period (date unrecorded). There was a somewhat higher than expected proportion of women who were within 8–16 days since their last period started (in their postmenstrual phase).

The clinician preferences for investigation echo the randomised options for the study. Almost all high-risk women would be given biopsy (92%) and visualisation (92%), the preference for moderate-risk women is very often biopsy (85%) but less often visualisation (78%), and for low-risk women the preference is ultrasound (48%), with no evaluation preferred for 24% and biopsy for 25%.

Other patient characteristics of potential relevance to outcomes are summarised in *Table 17*, separately for the risk groups.

There was a strong trend for women rating their general health as 'better than most women of the same age' to be in the higher risk groups, in particular the highest risk group (postmenopausal).

Women were asked how they would feel about the prospect of endometrial investigations in outpatients, theatre and under anaesthetic. There were trends across risk groups in each case, so that the lower the risk group the greater the proportion of women who answered 'no problem'. There was a non-significant trend for self-stated sensitivity to pain increasing as risk (of group) lowered. With regard to the statement that doctors should decide how to deal with bleeding problems, there was strongest agreement among the high-risk (oldest) women. There was no trend with respect to women wanting as much information as possible about their condition and wishing to be given choice regarding tests and treatments.

There was a consistent pattern for the women with lower risk to evidence worries about health, that is, to agree with the statement that bleeding problems are very worrying, and to answer that they worry very much about their health. However, more of the lower risk women dismissed as 'not very likely' the idea that their bleeding may be due to cancer, yet the only women answering 'very likely' to this statement were also in these premenopausal groups (four women).

Distributions by risk group for NEO Extraversion and Neuroticism factors and GHQ scales are shown in the boxplots in *Figures 3* and *4*.

TABLE 17 *Self-reported characteristics of potential relevance to outcomes, separately by risk group*

Characteristic	p^a	Risk group		
		High (n = 186, 184, 178)[b]	Moderate (n = 314, 315, 306)[b]	Low (n = 140, 139, 135)[b]
General health (%)				
Very good	0.475	41	37	34
Better than most women my age	0.001	33	19	11
'No problem' (%) with endometrial investigations …				
in outpatients	0.002	72	83	86
in theatre	0.115	39	43	51
requiring general anaesthetic	0.128	44	42	53
Sensitive to pain (% 'yes')	0.165	14	17	22
'Strongly agrees' (%) …				
doctor should decide treatment	0.262	23	18	17
she likes to be given choice about tests/treatment	0.366	35	41	39
she likes to be told as much as possible about condition	0.037	58	68	69
bleeding problems are worrying	0.079	24	24	34
Worry 'very much' re health (%)	0.019	4	6	12
'Not very likely' bleeding may be due to cancer (%)	0.001	59	77	78
NEO scores, median (quartiles)				
Neuroticism	0.003	18 (14, 23)	21 (15, 27)	21 (16, 28)
Extraversion	0.216	27 (23, 31)	28 (24, 31)	28 (25, 32)
Openness	0.155	26 (22, 30)	27 (23, 31)	25 (22, 30)
Agreeableness	0.041	34 (31, 36)	34 (31, 36)	33 (29, 36)
Conscientiousness	0.951	34 (31, 37)	35 (31, 38)	34 (31, 37)
GHQ scores, median (quartiles)				
A: somatic symptoms	0.001	5.0 (2.0, 8.0)	7.0 (4.0, 10.0)	6.0 (4.0, 11.0)
B: anxiety	0.001	4.5 (1.5, 8.0)	7.0 (3.0, 10.0)	6.0 (3.0, 10.0)
C: social dysfunction	0.197	7.0 (7.0, 8.0)	7.0 (7.0, 10.0)	7.0 (7.0, 9.0)
D: depression	0.044	0.0 (0.0,1.0)	0.0 (0.0, 3.0)	0.0 (0.0, 2.0)
Total	0.001	17 (12, 25)	21 (15, 30)	20 (15, 30)

[a] χ^2 test for comparison of percentages and Kruskal–Wallis test for comparison of personality scores.
[b] Numbers of women in risk group completing the Health questionnaire, the GHQ and the NEO, respectively. However, not all women completed all of the Health questionnaire. The number answering the different items ranged between 95 and 100% of n, except for the most poorly answered question 'bleeding due to cancer', which was answered by 80% in the low-risk group, 73% in the moderate-risk group and 65% in the high-risk group.

The high-risk group had lower NEO neuroticism scores (Kruskal–Wallis test, $p = 0.003$), and they also had lower scores on the GHQ scales for somatic symptoms, anxiety and depression, and for the total score. For social dysfunction the majority of women scored within a narrow range (7–10), but there were outliers with both high and low scores. One item on this scale asks 'Have you recently been able to enjoy your normal day-to-day activities?' A response of 'less so than usual' might be expected from a woman suffering excessively heavy menstrual periods, even if she was psychologically well. Such a response, together with responses of 'same as usual' for the other six items, would give a scale score of 8. For depression and to a lesser extent anxiety and somatic symptoms the GHQ scores were positively skewed, with a number of women in each risk group having high enough scores to be plotted as 'outliers'.

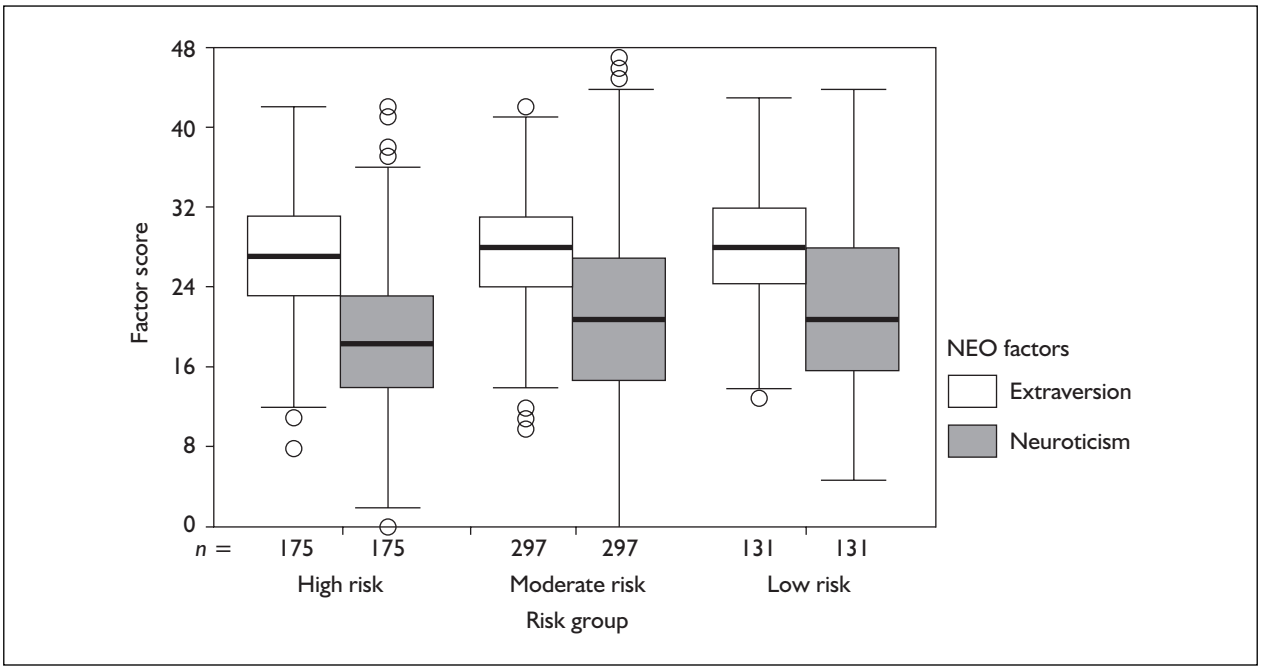


FIGURE 3 *Boxplot of NEO neuroticism and extraversion scores by risk group*

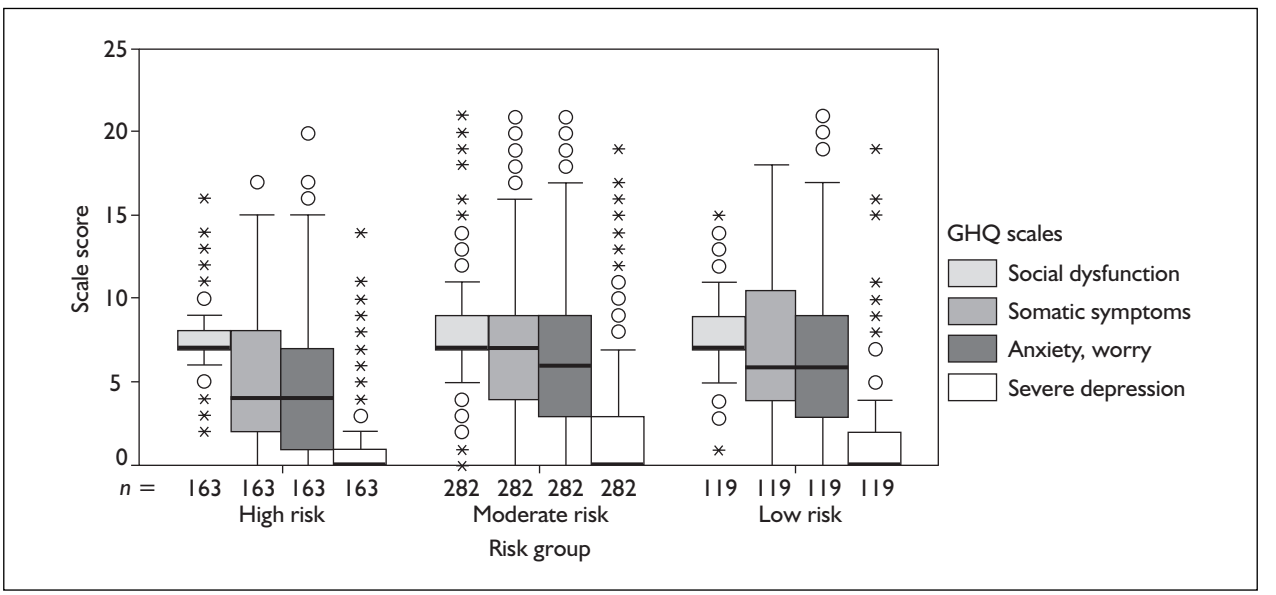


FIGURE 4 *Boxplot of GHQ scale scores by risk group*

Outcomes and estimation

Comparison of the three outpatient methods of evaluation

Successful completion of hysteroscopy, transvaginal ultrasound and biopsy

The outcome for investigations in the high-risk group is recorded in *Table 18*. There were non-obscured views for 79 women assigned to hysteroscopy and 87 women assigned to ultrasounds. In terms of biopsy, for Pipelle the numbers of women with an 'adequate' (or 'barely adequate') pathology sample were relatively low (50 and 36 women, respectively) compared with Tao brush (83 and 61 women, respectively).

As explained in Chapter 3, success rates have been calculated using different denominators, for example, ITT, as a percentage of the total number of women randomised to that option (N_t in *Table 18*), and also as a percentage of the number of women where the examination was potentially medically possible (N_m) or completed (N_c). These percentage success rates for high-risk women are presented in *Table 19*.

TABLE 18 *High-risk women: numbers by degree and 'success' of completion of investigations*

Total number of women (N_t)	**Randomised visualisation**	
Visualisation	**Hysteroscopy 100** ***n***	**Ultrasound 100** ***n***
Not possible for medical/equipment reason	13	0
Women for whom investigation is medically possible (N_m)	87	100
Not done for other reasons (e.g. DNA)	3	5
Women with investigation completed (N_c)	84	95
'Success': non-obscured view[a]	79	87
Biopsy[b]	**At hysteroscopy**	**Blind**
	P ***n*** **T** ***n***	**P** ***n*** **T** ***n***
Women with biopsy medically possible (N_m)	90 90	75 75
Women with sample taken (N_c)	89 89	75 75
'Success': adequate[c] pathology sample	50 83	36 61

The relevant numbers of women for analyses on the ITT basis are indicated by N_t, while for descriptive clinical summaries the numbers of women for whom the investigation was *medically possible* are indicated by N_m, and the numbers *actually completed* are given by N_c.
[a] For ultrasound, includes views noted as 'difficult'.
[b] All high-risk women were randomised to both Pipelle biopsy (P) and Tao brush biopsy (T).
[c] 'Adequate' or 'barely adequate' pathology sample.
DNA, did not attend.

TABLE 19 *High-risk women: success rates for visualisation and biopsy, by type of visualisation*

Denominator[a]	**Investigation type**	**Investigation outcome**		
		Successful visualisation (% success)	**Adequate Pipelle sample (% success)**	**Adequate Tao sample (% success)**
ITT: all in group (N_t)[a]	Hysteroscopy	79	50	83
	Ultrasound	87	36	61
	p (χ^2 test)	0.188	0.063	0.001
Those medically possible (N_m)[a]	Hysteroscopy	91	56	92
	Ultrasound	87	48	81
Investigations completed (N_c)[a]	Hysteroscopy	94	56	93
	Ultrasound	92	48	81

[a] See *Table 18*.

The different findings on an ITT basis, compared with completed investigations, can be seen by examining the visualisation success rates for hysteroscopy versus ultrasound. On completed investigations they are very similar, with 94% for hysteroscopy and 92% for ultrasound, whereas on the ITT basis they differ by a greater amount, with 79% for hysteroscopy and 87% for ultrasound. Statistical comparisons were undertaken only for ITT summaries. There was no significant difference in rates of successful visualisation between the two randomisation options, and only a borderline significant increase in successful Pipelle biopsy samples, if undertaken with hysteroscopy, compared with blind. There was, however, a significant increase in adequate Tao samples if taken at the time of hysteroscopy (83 versus 61%, $p < 0.001$).

For the moderate-risk group the numbers with various outcomes for biopsy and visualisations are recorded in *Table 20*.

Overall (out of 326 women in this moderate-risk group) there were non-obscured views using

TABLE 20 *Moderate-risk women: numbers by degree/'success' of completion of investigations*

	Moderate risk women			
Number of women N_t	**H 84**	**U 80**	**H + U 82**	**Neither 80**
Hysteroscopy (H)	*n*	*n*	*n*	*n*
Not possible for medical/equipment reason	10	–	7	–
Medically possible (N_m)	74	–	75	–
Not done for other reasons (e.g. DNA)	3	–	3	–
Completed (N_c)	71	–	72	–
Success: non-obscured view	64	–	63	–
Ultrasound (U)				
Not possible for medical/equipment reason	–	0	0	–
Medically possible (N_m)	–	80	82	–
Not done for other reasons (e.g. DNA)	–	3	7	–
Completed (N_c)	–	77	75	–
Success: non-obscured view[a]	–	73	71	–

Biopsy[b]	**With H**		**Blind**		**With H**		**Blind**	
	P	**T**	**P**	**T**	**P**	**T**	**P**	**T**
	n	***n***	***n***	***n***	***n***	***n***	***n***	***n***
Biopsy medically possible (N_m)	78	78	64	64	73	72	69	69
Sample taken (N_c)	76	75	64	64	71	70	69	69
Success: adequate[c] pathology sample	71	63	59	55	65	66	62	59

The relevant numbers of women for analyses on the ITT basis are indicated by N_t, while for descriptive clinical summaries the numbers of women for whom the investigation was *medically possible* are indicated by N_m, and the numbers *actually completed* are given by N_c.

[a] For ultrasound, includes views noted as 'difficult'.

[b] All moderate-risk women were randomised to both Pipelle biopsy (P) and Tao brush biopsy (T).

[c] 'Adequate' or 'barely adequate' pathology sample.

hysteroscopy for 127 women and using TVUS for 144, and there were adequate pathology samples for 257 women biopsied by Pipelle sampler and for 243 women biopsied by Tao brush. For these women, factorially randomised to hysteroscopy or ultrasound, in addition to both biopsies, the percentage success rates are presented in *Table 21*, separately for the three denominator types. There was a significantly higher rate of successful visualisation for ultrasound compared with hysteroscopy (89 versus 77%, $p = 0.005$), and no difference in the rate of successful Pipelle or Tao biopsy samples, if undertaken with hysteroscopy, compared with blind.

Comparing these moderate-risk women (*Table 21*) with high-risk women (*Table 19*), for each method of visualisation, there were non-obscured views for a similar percentage in the two risk groups. However, in terms of biopsy the percentage of adequate pathology samples was much higher in moderate-risk women, generally around 70–80% for both types of endometrial biopsy device, compared with high-risk women, where the success rate was 50% or less for Pipelle sampler and 60–80% for Tao brush. Since both moderate- and high-risk groups provide a comparison of hysteroscopy with ultrasound, it is possible to combine the data over these two risk groups to increase the power of the comparison. The data are summarised in *Table 22*, on an ITT basis only. The success rate for visualisation was significantly better for ultrasound (11 percentage points, $p = 0.002$), but adequate samples were more common with hysteroscopy, for both Pipelle (10 percentage points, $p = 0.03$) and particularly for Tao (12 percentage points, $p = 0.002$).

For the low-risk women the numbers with successful completion of investigations are presented in *Table 23*. With respect to visualisation, a similar picture was seen as for higher risk women, in that successful visualisations were more frequent in the women having ultrasound investigation (97% versus 65%, $\chi^2 = 8.7$, 1 df, $p = 0.003$). The subgroup sizes are too small to draw strong inferences for success of biopsy, but there is no observed difference in the proportions

TABLE 21 *Moderate-risk women: success rates for visualisation and biopsy, by type of visualisation*

Denominator[a]	Investigation type[a]	Investigation outcome		
		Successful visualisation (% success)	Adequate Pipelle sample (% success)	Adequate Tao sample (% success)
ITT: all in group (N_t)[a]	Hysteroscopy	77	82	78
	Ultrasound	89	76	71
	p (χ^2 test)	0.005	0.209	0.226
Those medically possible (N_m)[a]	Hysteroscopy	85	90	86
	Ultrasound	89	91	86
Investigations completed (N_c)[a]	Hysteroscopy	89	93	89
	Ultrasound	95	91	86

[a] For visualisation, the denominator for hysteroscopy is obtained (from Table 20) by summing the corresponding denominators for the two subgroups receiving hysteroscopy, and for ultrasound, the two subgroups receiving ultrasound. Therefore, those in the hysteroscopy + ultrasound subgroup contribute relevant but independent visualisation data to both the hysteroscopy and ultrasound summaries.
For biopsy the denominator for 'with hysteroscopy' is obtained by summing the corresponding denominators for the two subgroups receiving hysteroscopy, and for 'blind' the two subgroups receiving ultrasound and 'no visualisation' (from Table 20).

TABLE 22 *High- and moderate-risk women combined: success rates for visualisation and biopsy, by type of visualisation: ITT*

Investigation type	Randomised visualisation			
	Hysteroscopy (n = 266) (% success)[a]	Ultrasound (n = 260) (% success)[a]	p (χ^2 test)	Difference: H% – U% (%) (95% CI)
Successful visualisation	77	88	0.002	–11 (–17 to –4)
Adequate Pipelle sample	70	60	0.027	10 (1 to 18)
Adequate Tao sample	80	67	0.002	12 (5 to 20)

[a] Accumulated data from ITT analyses in Tables 19 and 21.

of adequate biopsy samples, by blind or not, or Pipelle versus Tao.

Not all women randomised to TVUS (n = 294 for both ITT and medically potentially possible) received this investigation. Overall, 67 (23%) received abdominal ultrasound only and for 23 (8%) the type of ultrasound was not specified. Factors influencing the decision not to undertake TVUS may have been patient related (e.g. time, aversion to vaginal scanning probe) or professional judgement, where an adequate view was obtained by the preliminary abdominal ultrasound or risk was judged to be very low, or both. However, of the 66 women who had abdominal ultrasound instead of TVUS and had a view recorded, 13 (20%) were reported as difficult view (n = 8) or no view (n = 5), so it is unlikely that for them the decision not to proceed to TVUS was operator choice. Furthermore, in the high-risk group, for whom endometrial thickness is a particularly important diagnostic measure, 11 women had only abdominal ultrasound, and for nearly half of these (n = 5) there was a difficult or failed view. It should be borne in mind, therefore, that the success rate for obtaining specifically a TVUS visualisation would be lower than that presented in *Tables 18–23*. For high-risk women the transvaginal visualisation success rate would be 76%, very similar to hysteroscopy in that group, and over all three groups it would be 66%, lower than for hysteroscopy (77%).

Abnormalities and conditions detected by hysteroscopy, ultrasound and biopsy

Hysteroscopy and ultrasound were compared to assess the number of abnormalities and conditions that each investigation picked up, as shown in *Table 24*. Hysteroscopy is more likely to pick up polyps, whereas ultrasound is more likely to detect fibroids. Hysteroscopy has the advantage of detecting abnormalities in the cervix and cervical canal. Increased endometrial thickness (> 4 mm) is an indication of risk of endometrial cancer.

TABLE 23 *Low-risk women: numbers of successful completion of investigations*

Number of women (N_t)	**Low-risk women**		
	H (+B) **31**	**U** **32**	**B** **32**
Visualisation			
Not possible for medical/equipment reason	5	0	–
Women for whom investigation is medically possible (N_m)	26	32	–
Not done for other reasons (e.g. DNA)	2	1	–
Women with investigation completed (N_c)	24	31	–
Success: non-obscured view[a]	20	31	–
Biopsy[b]	**With H**		**Blind**
	P **(n = 17)**	**T** **(n = 14)**	**P** **(n = 17)**	**T** **(n = 15)**
Women with biopsy medically possible (N_m)	13	13	–	14	13
Women with sample taken (N_c)	12	12	–	14	12
Success: adequate[c] pathology sample	10	12	–	14	11

The relevant numbers of women for analyses on the ITT basis are indicated by N_t, while numbers of women for whom the investigation was *medically possible* are indicated by N_m, and the numbers *actually completed* are given by N_c.
[a] For ultrasound includes views noted as 'difficult'.
[b] All low risk women were randomised to *either* Pipelle biopsy (P) or Tao brush biopsy (T).
[c] 'Adequate' or 'barely adequate' pathology sample.

TABLE 24 *All women: abnormalities detected by hysteroscopy and ultrasound separately by risk group (effective* n*)*

	Hysteroscopy	**Ultrasound**	**p (χ^2 test)**
High-risk women (n = 84, 95)			
Possible endometrial cancer	3 (3.6)	–	–
Endometrial thickness > 4 mm	–	34 (39.1)	
Endometrial/uterine polyp	17 (20.2)	4 (4.2)	0.002
Uterine fibroids	7 (8.3)	29 (31.5)	<0.001
Cervix suspicious	1 (1.2)	–	–
Cervical polyp	9 (10.7)	–	–
Moderate-risk women (n = 143[a], 152[a])			
Endometrial/uterine polyp	19 (13.3)	7 (4.6)	0.015
Uterine fibroids	31 (21.7)	59 (39.3)	0.002
Cervix suspicious	0 (0)	–	–
Cervical polyp	7 (4.9)	–	–
Low-risk women (n = 24, 31)			
Endometrial/uterine polyp	1 (4.2)	2 (6.5)	0.999
Uterine fibroids	1 (4.2)	6 (20.0)	0.189
Cervical polyp	1 (4.2)	–	–

Data are shown as n (%).
[a] In the moderate-risk group 73 women were tested by both methods and contribute results to both columns.

However, the majority of the cases of increased thickness detected with the ultrasound will be false positives rather than true cases of endometrial cancer.

Altogether, 73 of the women in the moderate-risk group were investigated by both hysteroscopy and ultrasound, and it was striking how discordant the two methods were for detection of polyps and fibroids. In no patient was uterine polyps detected by both methods, and of the 12 discordant cases, with polyps detected by one method only, the odds that detection was by hysteroscopy were three times that for ultrasound (p = 0.15). For detection of fibroids there were ten women where fibroids were detected by both methods. For the remaining

TABLE 25 *All women: abnormalities detected by visualisation: ITT*

Abnormality	Randomised visualisation			
	H *n* = 297 (%)[a]	U *n* = 294 (%)[a]	*p* (χ^2 test)	Difference: H% – U% (%) (95% CI)
Uterine polyps	12.5	4.4	0.001	8 (4 to 12)
Fibroids	13.5	32.0	0.001	–19 (–25 to –12)

[a] Cases accumulated are from Table 24, denominators are ITT, i.e. all randomised.

TABLE 26 *All women: findings for randomised study biopsy (one or both) separately by risk group: effective* n[a]

Abnormality/condition detected by one or both biopsy methods	Risk group					
	High *n* = 146		Moderate *n* = 272		Low *n* = 46	
Endometrial cancer	5	3.4	3	1.1	0	0
Hyperplasia (atypical)	0	0	2	0.7	0	0
Hyperplasia (non-atypical)	1	0.7	1	0.4	0	0
Hyperplasia (simple)	1	0.7	0	0	0	0
Atrophic endometrium	12	8.2	0	0	0	0
Inactive endometrium	106	72.6	19	7.0	1	2
Cyclic endometrium	26	17.8	213	78.3	43	91
Other	27	18.5	59	21.7	0	0

Data are shown as *n* (%). In high- and moderate-risk women two pathology samples were taken so the percentages can and do add up to >100%.
[a] Excludes cases where no adequate sample obtained and cases where pathology results missing (*n*=3, 1, 1 respectively).

discordant cases, with fibroids detected by one method only ($n = 26$), the odds that detection was by ultrasound were 2.7 times that for hysteroscopy ($p = 0.03$). Therefore, ultrasound was significantly better than hysteroscopy at detecting fibroids.

In the low-risk women the number of abnormalities and conditions detected by the hysteroscopy or ultrasound, or both, was very low, and although ultrasound was more likely to detect fibroids than hysteroscopy, the difference was not statistically significant.

The abnormality detection rates on an ITT basis all reduce by a factor of approximately 0.988, but for the high- and moderate-risk groups the significant differences between hysteroscopy and ultrasound persist. Considering the three groups combined on an ITT basis the comparison of hysteroscopy with ultrasound for the detection of polyps and fibroids is shown in *Table 25*. Overall, polyps were detected significantly more often by hysteroscopy (8 percentage points, $p < 0.001$) and fibroids significantly more often by ultrasound (19 percentage points, $p < 0.001$).

The aim of biopsy is to exclude premalignant or malignant changes. In low- and moderate-risk women cancer is very rare, but there are other histological diagnoses that can provide information helpful to management of their bleeding symptoms. *Table 26* summarises the diagnoses of endometrial cancer and hyperplasia in the three groups, if detected by one or other randomised study biopsy method. *Table 26* shows that eight endometrial cancers and one case of complex atypical hyperplasia (CAH) were detected. A further case of endometrial cancer was detected by further investigations (inpatient hysteroscopy and endometrial curettage), and a case of cervical cancer was detected in a woman referred with abnormal vaginal bleeding who at initial clinical examination was diagnosed with a cervical tumour. (After the finding of the tumour none of the randomised investigations was undertaken. Cervical biopsy confirmed a poorly differentiated squamous carcinoma.)

Therefore, in total, there were nine cases of endometrial cancer, one case of CAH and one case of cervical cancer. By risk group, endometrial cancer (including CAH) was detected in 3.0% of high-risk women randomised and in 1.2% of moderate-risk women. The 11 cancer cases are detailed in *Table 27*.

TABLE 27 *Details of randomised investigations and findings for the 11 women in whom cancer was diagnosed*

Randomisation	Risk group	Outcome for randomised study investigations			
		H	**T**	**P**	**U**
H+T+P	High	Done, NF	Grade 1 end. cancer	Inadequate sample	NA
	High	Done, NF	Grade 2 end. cancer	Grade 2 end. cancer	NA
	High	Possible cancer	Grade 1 end. cancer	Grade 1 end. cancer	NA
	Moderate	Done, NF	CAH	Grade 3 end. cancer	NA
H+T+P+U	Moderate	Done, NF	CAH	CAH	9.6 mm
P+T+U	High	NA	Grade 3 end. cancer	Inadequate sample	1.9 mm
	High	NA	Grade 1 end. cancer	Grade 1 end. cancer	30.0 mm
	High[a]	NA	Failed insertion	Failed insertion	8.4 mm
	High[b]	NA	Not undertaken	Not undertaken	DNA/unnecessary
	Moderate	NA	Grade 2 end. cancer	Grade 2 end. cancer	Difficult view
P+T	Moderate	NA	Grade 2 end. cancer	Grade 2 end. cancer	NA

[a] After randomised investigations were attempted, an inpatient hysteroscopy and endometrial curettage were completed with finding of 'suspicion of malignancy'. Total abdominal hysterectomy and bilateral salpingo-oophorectomy was performed and a grade 2 endometrial cancer diagnosed.
[b] A diagnosis of cervical tumour was made on initial vaginal examination before the randomised investigations were completed. Cervical biopsy revealed a 'poorly differentiated squamous carcinoma'.
NF, nothing found; NA, not applicable (not randomised to this investigation); end., endometrial.

TABLE 28(a) *Numbers of women randomised and having investigations, and with inconclusive results and/or no reference test results for endometrial cancer*

	Randomised N_t	**Test complete N_c**	**Index tests: 'inconclusive' results[a] *n* (% of N_c)**	**Reference test: no result[c] (i.e. not done/fail/no result) *n* (% of N_c)**
TVUS[a]	100	95	8 (8.4)	31 (32.6)
H	297	287	61[b] (21.3)	33 (11.5)
P[d]	543	528	175[b] (33.1)	55 (10.4)
T[d]	540	523	124[b] (23.7)	46 (8.8)

[a] Includes only postmenopausal women (test cut-off of > 4 mm indicative of cancer applies only to postmenopausal women).
[b] For index tests 'inconclusive' results are counted as negative results. 'Inconclusive results' include poor visualisations or inadequate samples but they also include, for biopsies and hysteroscopy, failed insertions (meaning no result is possible). For hysteroscopy they also include randomised tests that were not undertaken for reasons related to the nature of the test (e.g. unavailability of sterilised hysteroscope, patient too frail for procedure).
[c] For reference tests 'inconclusive' results are counted a no result.
[d] Excludes low-risk women randomised to biopsy alone (as they had no other randomised investigations which could have detected cancer).

TABLE 28(b) *Sensitivity, specificity, positive and negative predictive values for endometrial cancer, for all four outpatient methods of investigation*

	Sensitivity	**Specificity**	**PPV**	**NPV**
TVUS[a] (*n* = 64)	66.7 (20.8, 93.9)	55.7 (43.3, 67.5)	6.9 (1.9, 22.0)	97.1 (85.5, 99.5)
H[a] (*n* = 254)	20.0 (3.6, 62.4)	98.8 (96.5, 99.6)	25.0 (4.6, 69.9)	98.4 (96.0, 99.4)
P[a] (*n* = 473)	70.0 (39.7, 89.2)	100 (99.2, 100)	100 (64.6, 100)	99.4 (98.1, 99.8)
T[a] (*n* = 478)	90.0 (59.6, 98.2)	100 (99.2, 100)	100 (70.1, 100)	99.8 (98.8, 100)

Data are shown as % (95% CI, calculated using Wilson's method.[64])
[a] See *Table 28a*.

Table 28(a, b) summarises the numbers of tests and results for detection of endometrial cancer (nine cases plus one CAH), and report sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) with confidence intervals.

The small number of cases overall (because the main focus of the study was patient acceptability and economic evaluation) means that there are very wide confidence intervals for all sensitivities, in particular for ultrasound, because there were only three cases of endometrial cancer in postmenopausal women randomised to TVUS. For hysteroscopy and ultrasound the only women excluded from the calculation were those who failed to attend their appointment. For biopsy there were very few failures to attend, but women were also excluded if the test was 'not done for medical reasons', as this happened fairly often as a consequence of a preceding hysteroscopy that had made the woman distressed. That is, the 'medical reason' was not connected with the biopsy procedure itself.

Inconclusive index test results were included as negative results on the basis that no cancer was found by that test, whereas it might have been by some other test. Therefore, this strategy allows a more realistic assessment of sensitivity, in the sense of picking up true cancers that could be detected by the reference tests. However, this strategy will tend to overestimate specificity. If the inconclusive index tests were excluded for those with no cancer detected by reference tests, rather than being counted as negative results, then the specificity for the four diagnostic methods would be TVUS 52% (n = 85), hysteroscopy 98% (n = 226), Pipelle 100% (n = 353) and Tao 100% (n = 400).

Time taken for investigations

Relative time taken to complete an investigation was consistent across the three risk groups, with blind biopsies taking the least time (mean 3–5 minutes), hysteroscopy (with biopsy) an intermediate time (means 10–12 minutes) and ultrasound the most time (13–15 minutes) (*Table 29*). For blind biopsy and ultrasound the amount of time taken tended to decrease as risk decreased, which may be partly explained by the fact that low-risk women had only one biopsy, whereas in the other two groups all women were assigned two biopsies. For hysteroscopy the trend was in the reverse direction, with the hysteroscopy taking longer as risk decreased, this despite the fact that the low-risk women needed only one biopsy to be done. There was little saving in time if abdominal ultrasound only was undertaken, and the few women who had TVUS only tended to take as long or longer for their investigations.

Minor adverse events

All adverse events were minor. There were no adverse events at all for ultrasound, but for hysteroscopy and biopsy there were overall 12% and 9%, respectively (effective n). The adverse events for hysteroscopy and blind biopsy are detailed in *Table 30*, separately by risk group.

To undertake an ITT comparison it is necessary to count those women who did not have their investigation as having had an adverse event. *Table 31* summarises the percentages of adverse events by investigation and risk group, using this ITT basis. There was a significant difference between the three investigations, but this was mainly due to the low rate of adverse events with ultrasound [6% compared with 16% with hysteroscopy (plus

TABLE 29 *Duration of investigations separately for risk groups*

Risk group	*n*	Duration of investigation (minutes)				
		Mean (SE)	Median	Lower quartile	Upper quartile	Range
High						
U	90	15.3 (0.59)	15	13.0	15.0	5–50
H + B	93	9.7 (0.54)	10	5.0	10.0	2–35
Blind B	93	5.1 (0.38)	5	3	5	1–25
Moderate						
U	148	14.8 (0.47)	15	11.5	15.0	6–68
H + B	156	11.3 (0.49)	10	7.0	15.0	2–35
Blind B	146	4.2 (0.26)	4	2	5	1–20
Low						
U	30	13.1 (0.52)	15	10.0	15.0	9–21
H + B	28	11.8 (1.21)	10	8.5	15.0	2–30
Blind B	29	2.8 (0.40)	2	1	5	1–10

TABLE 30 *All women: minor adverse events with hysteroscopy and blind biopsy: effective* n

	Adverse events by risk group		
	High	**Moderate**	**Low**
Hysteroscopy (*n* = 94, 159, 28)			
Shock	1 (1.1)	5 (3.1)[1]	0 (0.0)
Possible perforation/wrong cavity	0 (0.0)	1 (0.6)[1]	0 (0.0)
Patient distress	8 (8.5)[1]	8 (5.0)[2]	7 (25.0)[1]
Trauma to cervix	1 (1.1)	0 (0.0)	0 (0.0)
Pain	0 (0.0)	2 (1.3)	0 (0.0)
Total	10 (10.6)	16 (10.1)	7 (25.0)
Blind Biopsy (*n* = 99, 159, 30)			
Shock	1 (1.0)[1]	0 (0.0)	0 (0.0)
Perforation	0 (0.0)	0 (0.0)	0 (0.0)
Patient distress	10 (10.1)[3]	9 (5.7)[7]	3 (10.0)[1]
Pain	0 (0.0)	1 (0.6)[1]	0 (0.0)
Feels faint/fainted	0 (0.0)	1 (0.6)	0 (0.0)
Shock and feels faint/fainted	0 (0.0)	1 (0.6)	0 (0.0)
Total	11 (11.1)	12 (7.5)	3 (10.0)

Data are shown as *n* (%).
The superscript figures (e.g.[1]) represent the number of women with hysteroscopy and/or biopsy not completed because of the adverse event. So the majority of high-risk women had the investigation despite an adverse event occurring (although the biopsies were not necessarily completed if an adverse event occurred for hysteroscopy), but this was not the case for moderate-risk women.

TABLE 31 *All women: minor adverse events (%) by investigation using ITT analysis (with adverse events imputed for investigations not completed and/or where adverse event data missing)*

% Adverse events (ITT)	**Risk group**			
	High (*n* = 100)	**Moderate (*n* = 160–166)**	**Low (*n* = 31–32)**	**All women (*n* = 292–297)**
Ultrasound (U)[a]	7	6	3	6
Hysteroscopy (with biopsy) (H)	16	14	32	16
Blind biopsy (B)	12	8	16	10
Comparisons, *p* (χ^2 test)				
Comparing U, H and B	0.139	0.046	0.008	<0.001
H vs B	0.541	0.141	0.210	0.036

[a] Adverse event data missing for two high-risk patients who had ultrasound completed.

biopsy) and 10% with blind biopsy, $p < 0.001$]. Nevertheless, the difference in rate of adverse events between hysteroscopy (with biopsy) and blind biopsy shows a weak trend for hysteroscopy to be associated with more adverse events (overall, 16% versus 10%, $p = 0.036$).

Comparing the three investigation methods in terms of adverse events as a percentage of completed investigations, there was a significant difference in rate of adverse events, overall and within each risk group (*Table 32*). However, there was no significant difference in adverse events between hysteroscopy and biopsy. For investigations undertaken at the same time (blind biopsy by both Pipelle and Tao samplers, or hysteroscopy with biopsy), the data collection regarding adverse events allowed recording of such events only for the investigation package overall, and did not attribute the adverse event to a particular part of the evaluation. However, as can be seen by perusing the list in *Table 30*, the most common adverse events were typically reactions to insertion, which would usually become apparent at the first procedure to be undertaken. For blind biopsy it would be the first sampler to be used, which was randomised to be 50% Pipelle and 50% Tao, and for hysteroscopy with biopsy this would be the hysteroscopy.

TABLE 32 *All women: minor adverse events (%) by investigations attempted and/or completed, separately for risk groups and overall*

% Adverse events (of completed investigations)	Risk group			
	High (*n* = 93–99)	**Moderate (*n* = 152–159)**	**Low (*n* = 28–31)**	**All women (*n* = 276–288)**
Ultrasound (U)	0	0	0	0
Hysteroscopy (with biopsy) (H)	11	10	25	12
Blind biopsy (B)	11	8	10	9
Comparisons, p (χ^2 test)				
Comparing U, H and B	0.004	0.001	0.010	0.001
H vs B	0.999	0.553	0.245	0.355

It might be expected that insertion of the thicker hysteroscope would be more difficult than insertion of a sampler, and so may be more likely to result in an adverse event. However, this only looks as if it may have been the case in the very small low-risk group, for which estimated proportions are very unreliable. Even if insertion failed for hysteroscopy, it was sometimes possible to undertake biopsy, provided the woman was not too upset. For the 20 women overall for whom hysteroscopy failed, biopsy samples were nevertheless obtained for 5 (25%) of them at the same appointment.

Women's (short-term) report about clinic visit and acceptability of procedure

Acceptability in the short term was assessed by means of:

- rating the unpleasantness (or not) of the investigation
- reporting postinvestigation on the after-effects, abdominal discomfort and bleeding
- reporting their feelings about the clinic visit (whether they are glad to have had the investigation, how reassured they are, and whether they would have liked more investigation)
- ascertaining each woman's subjective judgement as to how worthwhile the clinic visit has been.

Other feelings about the clinic visit were also assessed, but they are not the main outcomes, serving rather to provide a broader understanding of the women's experiences of outpatient endometrial investigation.

Finding investigation(s) unpleasant

Table 33 shows separately by risk group the numbers reporting on their investigations and the numbers rating each markedly unpleasant. Percentages calculated on an ITT basis are presented by risk group and overall, and compared across the three investigation methods. For clinical interest, the percentages calculated using the effective n, the number of investigations completed and reported, are also presented. The prevalences of unpleasant ratings on this basis were all lower than by ITT, but the pattern of responding was unchanged.

Rating investigations as unpleasant was on the whole most prevalent among the low-risk group, but these rates should be interpreted with caution, given the small number of women involved. Only two women overall found the ultrasound procedure markedly unpleasant, so there was a consistent pattern across all three risk groups that ultrasound was least often rated unpleasant (11% ITT in the high- and moderate-risk groups combined). In the high- and moderate-risk women, hysteroscopy and biopsy were more often rated as markedly unpleasant (in high- and moderate-risk women combined, 26% and 29% ITT, respectively, for the two methods). The differences in ITT percentages between the three methods were statistically significant in the high- and moderate-risk groups, and borderline in the low-risk group ($p < 0.001$, $p < 0.001$ and $p = 0.053$, respectively). *Figure 5* illustrates the corresponding data using effective n for calculation of percentages. While there was little difference between hysteroscopy and biopsy percentages in the high- and moderate-risk groups, in both groups ultrasound was much less often rated as unpleasant.

Ratings for women reporting on investigations that could not be completed were compared with the rest completing the same investigation. Overall, 15% of women with successfully completed hysteroscopies described the investigation as markedly unpleasant, compared with 30% of women with unsuccessful hysteroscopies. This difference was not significant (χ^2 test, $p = 0.157$). There was, however, a significant difference in the percentage of 'markedly unpleasant' ratings between women with

TABLE 33 *Finding investigation 'markedly unpleasant', by investigation method and risk group*

Risk group	Investigation method			
	H +B (*n*)	**U (*n*)**	**B (*n*)**	***p*** **(χ^2 test)**
High (N_t =100, 100, 200)				
At least one biopsy unsuccessful	–	–	36	
Question answered (N_e[a])	83	90	161	
'Markedly unpleasant' (observed, O_u)	11	1	27	
% Unpleasant (ITT[b])	28%	11%	33%	0.001
% Unpleasant (effective *n*[c])	13%	1%	17%	
Moderate (N_t =166, 162, 326)				
At least one biopsy unsuccessful	–	–	51	
Question answered (N_e[a])	149	147	296	
'Markedly unpleasant' (observed, O_u)	23	1	54	
% Unpleasant (ITT[b])	24%	10%	26%	0.001
% Unpleasant (effective *n*[c])	15%	1%	18%	
Low (N_t =31, 32, 63)				
At least one biopsy unsuccessful	–	–	13	
Question answered (N_e[a])	26	27	52	
'Markedly unpleasant' (observed, O_u)	8	0	12	
% Unpleasant (ITT[b])	42%	16%	37%	0.053
% Unpleasant (effective *n*[c])	31%	0%	23%	
Overall (N_t =297, 294, 589)				
% Unpleasant (ITT[b])	27%	11%	29%	0.001

[a] Some women with unsuccessful investigations nevertheless completed the report and the 'unpleasantness' question, but the majority of women with unsuccessful hysteroscopies were not given the hysteroscopy report to complete as the investigation was not even attempted.
[b] Those who did not have investigation/did not answer the unpleasantness question are imputed to have found the investigation unpleasant, so for this percentage the numerator is O_u+ [N_t – N_e] and the denominator is N_t.
[c] For this percentage the numerator is O_u and the denominator is N_e.
O_u, number of women with outcome present.

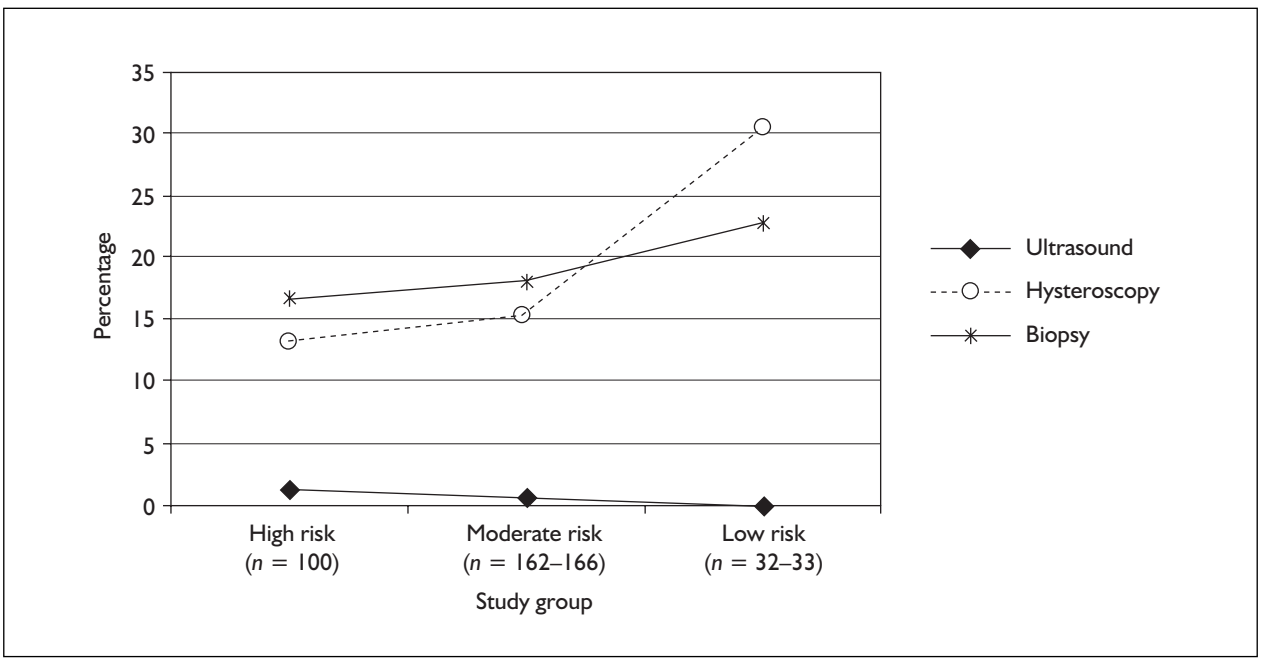


FIGURE 5 *Percentages of investigations rated as 'markedly unpleasant': effective* n

all biopsies completed successfully and those with at least one unsuccessful biopsy (17% versus 33%, χ^2 test, $p = 0.006$).

Abdominal discomfort with or after, and bleeding after, the investigation

Women were also asked immediately after biopsy to record on their investigation report the occurrence and severity of cramps or discomfort at the time of the biopsy. All high-risk and moderate-risk women were randomised to biopsy, but not all of them answered the question (163/200 high-risk and 298/326 moderate-risk women). Approximately one-third of high- and moderate-risk women reported cramps or discomfort (33% and 35%, respectively). The percentage was higher in the low-risk women randomised to biopsy, almost one-half (49%, $n = 63$ randomised, 51 answered the question). The percentage of women experiencing cramps or discomfort varied significantly depending on whether all the biopsies were successfully completed, or one (or both) of the biopsies had failed (34% versus 50%, χ^2 test, $p = 0.019$; $n = 452$ and 60 answering the question, respectively).

In the clinic review questionnaire completed at home on the day after the last investigation, women were asked whether they had suffered from cramps, bleeding or discomfort at home after their clinic visit(s). The questionnaire asked for an answer overall for the clinic investigations, as where there had been multiple investigations (e.g. TVUS and hysteroscopy plus biopsy) it would be impossible to attribute any after-effects to one or other investigation. A single abdominal discomfort variable has been created from the two after-effects of cramps and discomfort, being for each woman the more severe of the two ratings given. The response choice was not at all, hardly any, some, a lot or severe. The latter two response categories were combined and are reported here as a binary outcome. As this is felt to be a matter of more interest to patient management, the percentages of women rating these after-effects 'a lot' or 'severe' are calculated on the basis of effective *n* (those reporting on after-effects). *Table 34* presents the percentages separately by group and randomisation option. In the postmenopausal women (high risk) postinvestigation bleeding was infrequent, and abdominal discomfort seemed less prevalent than in the premenopausal women. Comparing across

TABLE 34 *Percentage of women experiencing abdominal discomfort and/or bleeding at home after their clinic visit, separately by risk group: effective* n

Risk group	*p* (χ^2 test)	Percentage 'a lot' or 'severe'[a]			
		Randomisation (all have B)			
High Statements	**H vs U**	**H (*n* = 87)**		**U (*n* = 87)**	
Abdominal discomfort	0.158	21		12	
Bleeding	0.774	7		5	
		Randomisation (all have B)			
Moderate[b] Statements	**H vs No, U vs No**	**H (*n* = 144)**	**No H (*n* = 136)**	**U (*n* = 135)**	**No U (*n* = 145)**
Abdominal discomfort	0.418, 0.434	31.5	26.3	31.6	26.6
Bleeding	0.025, 0.445	21.5	10.8	14.3	18.4
		Randomised investigation			
Low Statements	**Among four options**	**H+B (*n* = 23)**	**B (*n* = 25)**	**U (*n* = 21)**	**None (*n* = 45)**
Abdominal discomfort	0.927	30	24	24	23
Bleeding	0.928	14	8	10	9

[a] Not all women who completed the form answered every question.
[b] For moderate-risk women, where the options were factorially randomised, and all women have biopsy, data are shown separately for a comparison of hysteroscopy versus no hysteroscopy, and ultrasound versus no ultrasound. In this section there were 67 women who had both investigations and therefore contribute data to the first and third columns, and 68 women who had neither visualisation, and therefore contribute data to the second and fourth columns.

randomisations, there was a non-significant trend in high-risk women for hysteroscopy to be followed by more discomfort than ultrasound, and in the moderate-risk women there was significantly more bleeding after hysteroscopy ($p = 0.025$). The numbers in the low-risk group are too small to interpret confidently, but the subgroup of women having hysteroscopy (with biopsy) reported the most discomfort (30%) postinvestigation and the most bleeding (14%).

Women were given the opportunity to report on other after-effects, which they specified in the space provided, using the same response options as for bleeding and discomfort. Sixteen women noted that they had suffered 'a lot' with backache ($n = 2$), nausea/vomiting ($n = 2$), discharge ($n = 2$), pain ($n = 3$), exhaustion ($n = 4$), emotional reaction ($n = 1$), tender stomach ($n = 1$) and sense of shock ($n = 1$). A further seven women each rated as 'severe' one of the following: backache, pain, exhaustion, emotions, anger, shaky and stressed. Twelve women suffered other effects but did not specify the level of their discomfort: backache ($n = 2$), nausea/vomiting ($n = 3$), urinary infection ($n = 1$), cold ($n = 1$), exhaustion ($n = 3$), periods started ($n = 1$) and dizzy ($n = 1$).

Feelings about the clinic visit

Regarding feelings about the clinic visit, ascertained on the day after the last randomised investigation,

TABLE 35 *Percentage of women expressing agreement with the statements about their feelings about the clinic visit(s), separately by risk group: effective* n

Risk group	**Percentage expressing agreement[a]**			
High (*n* = 174 women completing questionnaire)	**Randomisation (all have B)**			
Statements	**H (*n* = 87)**		**U (*n* = 87)**	
I am more worried than before the clinic attendance	11		11	
I do not really understand what the doctor told me about my bleeding	10		16	
I wish I had not bothered	4		2	
I would have liked more investigation of my bleeding problem	13		15	
I feel reassured by the visit	88		87	
I am glad I had the investigation	92		95	
Moderate (*n* = 280[a,b] women completing questionnaire)	**Randomisation (all have B)**			
Statements	**H (*n* = 144)**	**no H (*n* = 136)**	**U (*n* = 135)**	**no U (*n* = 145)**
I am more worried than before the clinic attendance	12.9	12.8	9.8	15.6
I do not really understand what the doctor told me about my bleeding	15.1	13.7	15.7	13.3
I wish I had not bothered	3.6	5.2	6.8	2.1
I would have liked more investigation of my bleeding problem	18.3	25.6	21.7	22.0
I feel reassured by the visit	84.4	90.4	90.0	85.2
I am glad I had the investigation	90.6	98.5	94.0	95.0
Low (*n* = 114 women completing questionnaire)	**Randomised investigation**			
Statements	**H+B (*n* = 23)**	**B (*n* = 25)**	**U (*n* = 21)**	**None (*n* = 45)**
I am more worried than before the clinic attendance	22	12	24	11
I do not really understand what the doctor told me about my bleeding	17	12	14	16
I wish I had not bothered	4	4	5	9
I would have liked more investigation of my bleeding problem	32	40	43	36
I feel reassured by the visit	91	88	71	82
I am glad I had the investigation	95	100	95	92[c]

[a] Not all women who completed the form answered every question.
[b] For moderate-risk women, data are shown separately for a comparison of hysteroscopy versus no hysteroscopy, and ultrasound versus no ultrasound.
[c] Only eight women out of the 44 who answered this question gave the response 'not applicable', so they may be referring to the clinical examination or to other non-outpatient methods of endometrial evaluation (OMEE) investigations.

the responses agreeing with the six statements are presented in *Table 35* where, as for after-effects, percentages are calculated on the effective *n*, those completing the questionnaire. Over 90% of women were at that time 'glad that they had had the investigations' and over 80% of women reported that they were reassured by the clinic visit. It is noteworthy that overall 15% of low-risk women 'did not really understand' what they were told about their bleeding problem, and similarly for 14% each of moderate- and high-risk women.

There were no marked differences between the methods, but the moderate-risk women who had had hysteroscopy were less glad (91% versus 99%, $p = 0.01$) about having had their investigation than those who had not had hysteroscopy were about their investigation (which would be biopsy only or ultrasound and biopsy). It is noteworthy that in all risk groups the women who had hysteroscopy were least likely to agree that they would have liked more investigation of their problem (13%, 18.3% and 32% agreeing in the high-, moderate- and low-risk groups, respectively). In comparison, among women who had biopsy only, 28% of moderate-risk women and 40% of low-risk women 'would have liked further investigation'.

There are small numbers in the low-risk group, so results must be interpreted with caution, but across the randomisation options there was a consistent wish for more investigation (37% overall); however, this was not highest in those randomised to no investigation. Nearly all of the low-risk women who were randomised to no investigations were 'glad that they had the investigations', so must have been referring to non-study investigations or considering the clinic examination as an investigation. They were also less likely to be 'more worried than before the clinic attendance' (11%) than other low-risk women who had at least one investigation.

How worthwhile women considered the clinic visit
Calculating the percentages on an ITT basis, the majority of high- and moderate-risk women felt the clinic visit was very or extremely worthwhile, regardless of randomisation option (*Table 36*). Ignoring randomisation options, the percentage judging the clinic visit as very or extremely worthwhile was highest in the high-risk group, intermediate in the moderate-risk group, and lowest in the low-risk group (77%, 63% and 46%, respectively, χ^2 test, $p < 0.0001$), but there were minimal differences between investigation options within groups.

For clinical interest, *Figure 6* shows the distribution of ratings across randomisation options within risk groups. Ratings not shown are 'not very' and 'not at all'. Therefore, the overall height of columns shows the percentage rating visit 'quite worthwhile' or better. It can be seen that there were minimal differences across groups and investigations in overall percentages, which were 95% or more even in the low-risk group. It was only in respect of ratings of 'very/extremely worthwhile' that differences can be seen between the risk groups, but there were no marked differences between randomisation options.

Women's self-report of outcome and health (at 10 and 24 months postevaluation)
To simplify reporting, the follow-up questionnaire data will be reported for the high- and moderate-risk groups only. The subgroup sizes in the low risk group are too small to make follow-up data summaries useful.

For the 10-month follow-up, *Table 37* summarises responses on the main items in the questionnaire, separately for high- and moderate-risk women. Ordinal variables were recategorised as binary, and data are reported as percentages of effective *n*. This summary is ITT in the sense that all women randomised to a specific investigation were surveyed, and their data included if they responded, but data for those not responding were not imputed.

In the moderate-risk group there were minimal differences across investigations for the persistence of symptoms (rates of 46–55% depending on investigation) or failure to cure the problem (26–27%), but in the high-risk group women

TABLE 36 *Percentages judging clinic visit as very or extremely worthwhile, by risk group: ITT*

Risk group	High		Moderate				Low			
Investigation	H+B	U+B	H+B	U+B	B	H+U+B	H+B	B	U	None
n	100	100	84	80	80	82	31	32	32	62
% Rating worthwhile	80	73	67	64	61	62	42	56	41	47

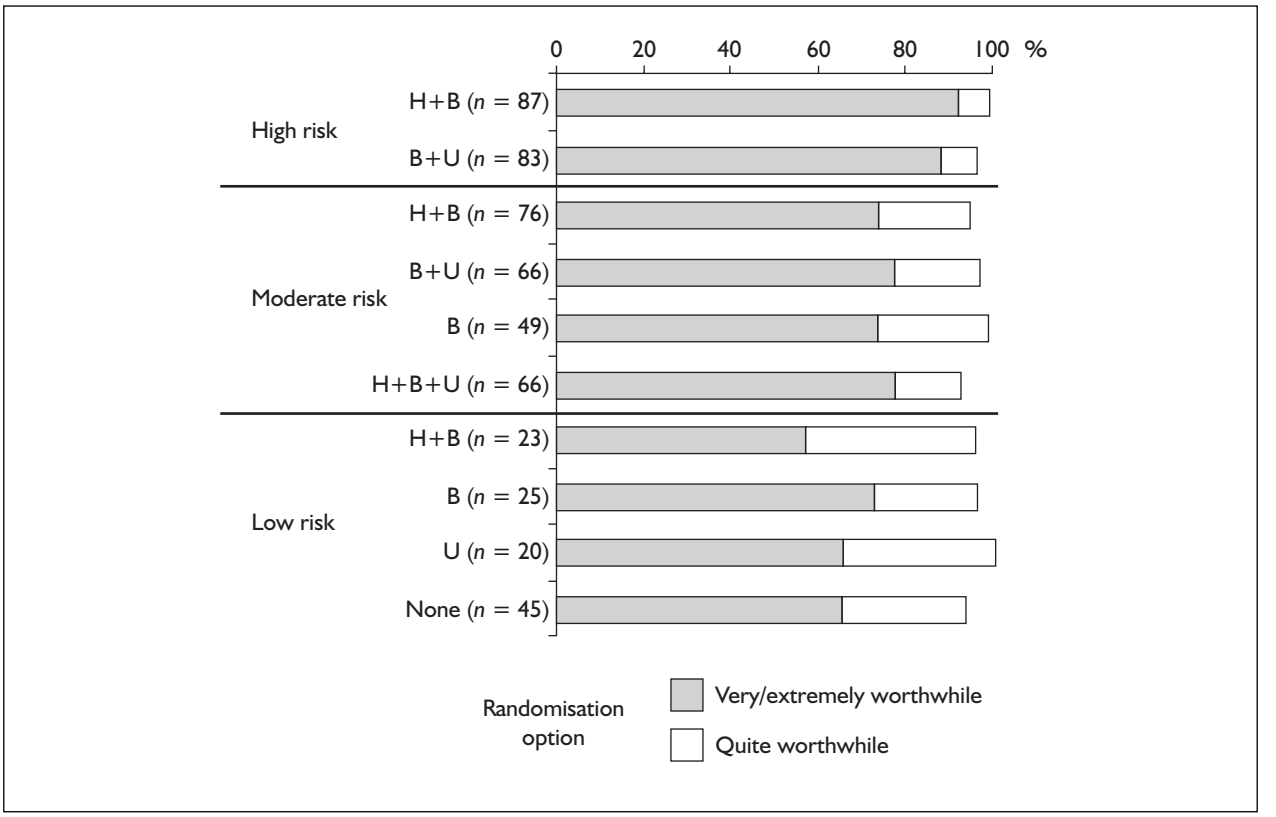


FIGURE 6 *Percentages reporting clinic visit as quite or very/extremely worthwhile, separately by randomisation option and risk group: effective* n

investigated by hysteroscopy tended to report better outcomes than those investigated by ultrasound (persisting symptoms, 15% versus 17%; failure to cure, 10% versus 14%). In the high-risk group there were advantages for hysteroscopy compared with ultrasound, in feeling satisfied with care and reassured by clinic attendance (11 percentage points in each case). This trend was also evident in the moderate-risk group, with a noticeable effect for hysteroscopy (versus not) for these two responses (15 and 8 percentage points difference, respectively), and for judging attending the clinic as very or extremely worthwhile (13 percentage points). However, in the moderate-risk group there was a similar, albeit slightly weaker, effect for these items for ultrasound (versus not), showing an increase of 6–8 percentage points.

Given the factorial assignment of visualisations, the similar advantage observed for both hysteroscopy and ultrasound (versus not) must therefore arise as a result of a more negative view in those who have neither hysteroscopy nor ultrasound, the biopsy-only subgroup. In the moderate-risk group this trend for biopsy-only women to be relatively dissatisfied was also evident in terms of wishing to have had more investigation, with 42% of women randomised solely to biopsy agreeing that this statement was fairly or very true.

Of women initially randomised, the proportions with missing responses for 10-month follow-up were 19% in the high-risk group, 20% in the moderate-risk group and 37% in the low-risk group. A sensitivity analysis for missing data would involve assuming the worst and best possible outcomes for the missing responses, so the interpretation of *Table 37* must be cautious. *Table 38* provides a similar summary for the data obtained by the 24-month follow-up questionnaire.

It can be seen that in the high-risk group the differences between hysteroscopy and ultrasound noted at 10 months had largely disappeared by 24 months. However, more women investigated by hysteroscopy than ultrasound were glad that they had had the investigation (90% versus 82%), and a higher proportion of women investigated by ultrasound responded they would have liked more investigation (19% versus 11%).

In the moderate-risk group a similar pattern of responding was seen as for the 10-month follow-up, with both comparisons (hysteroscopy versus not and ultrasound versus not) showing more

TABLE 37 *Ten-month follow-up: percentages for selected responses by women regarding clinic attendance, separately by high- and moderate-risk group: effective* n

Risk group	**% of effective *n***			
	Randomisation (all have B)			
High (*n* = 162 women completing questionnaire) Statements	**H (*n* = 83)**		**U (*n* = 79)**	
Symptoms much improved (%)	78		80	
Satisfied with care (% very true)	86		75	
Reassured by clinic attendance (% very true)	81		70	
Glad attended clinic (% very true)	86		87	
Worthwhile attending clinic (% 'very' or 'extremely')	90		86	
Symptoms persisting (% yes)	15		17	
Attendance failed to cure problem (% very true)	10		14	
Would have liked more investigation (% fairly/very true)	9		11	
	Randomisation (all have B)			
Moderate (*n* = 261[a,b] women completing questionnaire) Statements	**H (*n* = 136)**	**No H (*n* = 125)**	**U (*n* = 133)**	**No U (*n* = 128)**
Symptoms much improved (%)	42	38	38	42
Satisfied with care (% very true)	65	50	62	54
Reassured by clinic attendance (% very true)	64	52	61	55
Glad attended clinic (% very true)	71	65	70	67
Worthwhile attending clinic (% 'very' or 'extremely')	75	62	73	65
Symptoms persisting (% yes)	49	53	53	49
Attendance failed to cure problem (% very true)	27	26	27	26
Would have liked more investigation (% fairly/very true)	20	35	22	32

[a] Not all women who completed the form answered every question.
[b] For moderate-risk women, data are shown separately for a comparison of hysteroscopy versus no hysteroscopy, and ultrasound versus no ultrasound.

favourable responses for the visualisation in question, indicating that the underlying pattern is that the most favourable responses were generally for the option involving ultrasound and hysteroscopy (with biopsy), and the least favourable responses were in the subgroup receiving biopsy only. For example, 38% of women in the biopsy-only subgroup stated at 2-year follow-up that they would have liked more investigation.

Apart from the comparison across randomisation options, *Tables 37 and 38* also show clinic outcome of women referred for postmenopausal bleeding (high-risk group) compared with premenopausal women referred with menstrual bleeding problems (moderate-risk group). While the outcome for postmenopausal women was generally very good, fewer than half of the moderate-risk women (with menstrual bleeding problems) had much improved symptoms by 10 months (*Table 37*). Relative to postmenopausal women, almost three times as many moderate-risk women had persisting symptoms, and twice as many stated the clinic attendance had failed to cure the problem and that they would have liked more investigation of their problem. At the 2-year follow-up the same pattern was evident, but less strong, partly because the responses for the high-risk women were slightly less favourable than they had been at 10 months.

Comparison of biopsy methods: Tao brush sampling versus Pipelle sampling

Comparison of successful completion rates

In high-risk women, there was no difference in the proportion of successfully completed biopsies between the Tao brush and Pipelle. All the occasions where the Tao brush was not undertaken, the Pipelle was not undertaken, and vice versa. Of all the 200 women randomised to both biopsies, four (2.0%) women did not have the biopsies for medical reasons, 31 (15.5%) women's biopsies were abandoned on account of failed insertion and one (0.5%) woman did not attend her biopsy appointment.

For the 326 moderate-risk women, all randomised to both biopsies, 275 (84.4%) had both biopsies

TABLE 38 *Two-year (24 months) follow-up: percentages for selected responses by women regarding clinic attendance, separately by high- and moderate-risk group: effective* n

Risk group	**% of effective *n***			
	Randomisation (all have B)			
High (*n* = 127 women completing questionnaire) Statements	**H (*n* = 59)**		**U (*n* = 68)**	
Symptoms much improved (%)	79		81	
Satisfied with care (% very true)	86		86	
Reassured by clinic attendance (% very true)	81		82	
Glad attended clinic (% very true)	90		82	
Worthwhile attending clinic (% 'very' or 'extremely')	88		84	
Symptoms persisting (% yes)	15		17	
Attendance failed to cure problem (% very true)	16		18	
Would have liked more investigation (% fairly/very true)	11		19	
	Randomisation (all have B)			
Moderate (*n* = 182[a,b] women completing questionnaire) Statements	**H (*n* = 94)**	**No H (*n* = 88)**	**U (*n* = 92)**	**No U (*n* = 90)**
Symptoms much improved (%)	60	55	61	53
Satisfied with care (% very true)	73	53	67	60
Reassured by clinic attendance (% very true)	57	49	61	46
Glad attended clinic (% very true)	73	61	74	61
Worthwhile attending clinic (% 'very' or 'extremely')	71	62	68	65
Symptoms persisting (% yes)	42	48	44	46
Attendance failed to cure problem (% very true)	27	28	29	26
Would have liked more investigation (% fairly/very true)	16	31	17	29

[a] Not all women who completed the form answered every question.
[b] For moderate-risk women, data are shown separately for a comparison of hysteroscopy versus no hysteroscopy, and ultrasound versus no ultrasound.

TABLE 39 *High-risk women: percentage of acceptable samples in terms of quality for analysis for Tao brush and Pipelle samplers: effective* n

High-risk women		**N_e[a]**	**% of samples adequate or barely adequate**			
			Pipelle	**Tao brush**	**Difference (T–P) (95% CI)**	***p*-value[b]**
Hysteroscopy with biopsy	Brush first	37	56.8	97.3	40.5 (20.0 to 61.1)	<0.001
	Pipelle first	44	50.0	90.9	40.9 (20.0 to 61.8)	0.001
	Overall	89	56.2	93.3	37.1 (23.7 to 50.5)	0.001
Blind biopsy	Brush first	38	36.8	81.6	44.7 (23.5 to 66.0)	0.001
	Pipelle first	35	60.0	80.0	20.0 (–0.2 to 40.2)	0.096
	Overall	75	48.0	81.3	33.3 (18.8 to 47.9)	0.001
All high-risk women		164	52.4	87.8	35.4 (25.5 to 45.2)	0.001

[a] Women were included only if both Tao brush biopsy and Pipelle biopsy were successfully completed. Where the randomisation order was not specified or was uncertain those women were excluded from the subanalyses examining order.
[b] McNemar's paired comparison of Pipelle versus Tao within this subgroup.

successfully completed, failed insertion occurred for both biopsy methods for 37 women (11.3%) biopsies, and for six women neither biopsy was attempted, three for technical reasons and three for medical reasons. There were occasions where the Pipelle biopsy could be successfully completed but not the Tao brush biopsy, and vice versa. The Pipelle sampler was successful for five women where the Tao brush failed for medical reasons ($n = 3$) or failed insertion ($n = 2$), whereas the

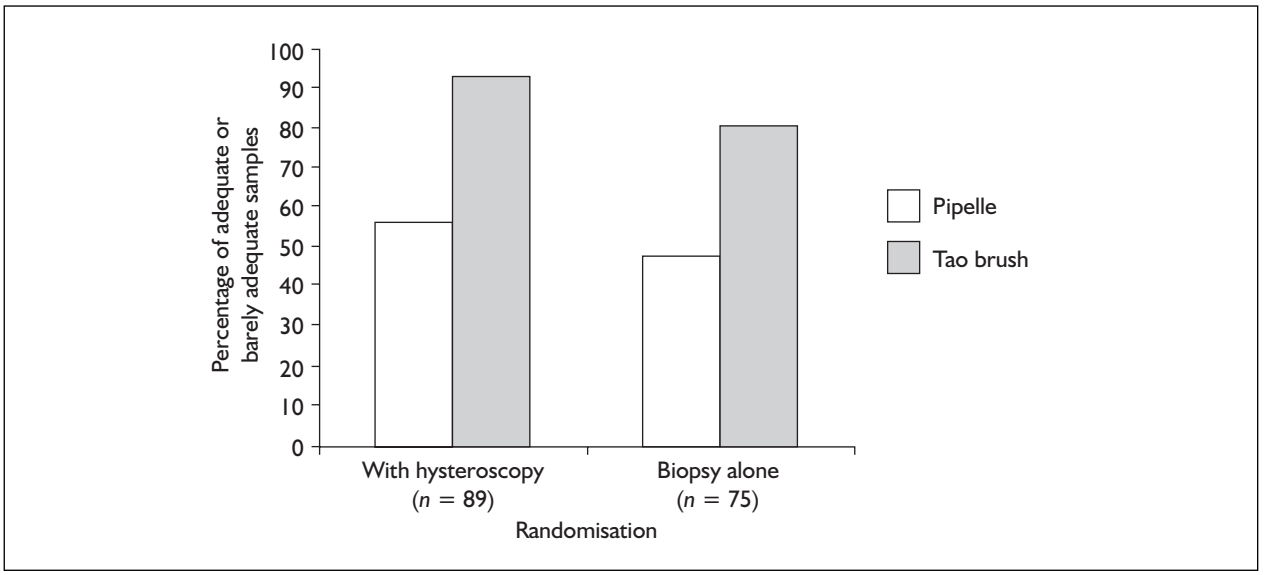


FIGURE 7 *High-risk women: percentage of acceptable samples for Tao brush and Pipelle biopsies: effective* n

TABLE 40 *Moderate-risk women: percentage of acceptable samples in terms of quality for analysis for Tao brush and Pipelle samplers: effective* n

Moderate-risk women		N_e[a]	% of samples adequate or barely adequate			
			Pipelle	Tao brush	Difference (P–T) (95% CI)	*p*-value[b]
Hysteroscopy with biopsy	Brush first	75	96.0	92.0	4.0 (–2.9 to 10.9)	0.450
	Pipelle first	66	89.4	84.8	4.5 (–6.2 to 15.3)	0.579
	Overall	147	92.5	87.8	4.8 (–1.6 to 11.2)	0.211
Blind biopsy	Brush first	64	93.8	82.8	10.9 (1.8 to 20.1)	0.046
	Pipelle first	65	86.2	84.6	1.5 (–9.3 to 12.4)	0.999
	Overall	136	89.0	83.8	5.1 (–1.8 to 12.1)	0.211
All moderate-risk women		283	90.8	85.9	4.9 (0.2 to 9.6)	0.055

[a] Women were included only if either Tao brush biopsy or Pipelle biopsy were successfully completed. Where the randomisation order was not specified or was uncertain those women were excluded from the subanalyses examining order.
[b] McNemar's paired comparison.

Tao brush was successful for three women where the Pipelle sampler failed for medical reasons ($n = 1$) or failed insertion ($n = 2$).

Comparison of adequate sample rates

In high-risk women with both biopsies successfully completed biopsies were compared for Pipelle sampler against Tao brush, in terms of obtaining an acceptable (adequate or barely adequate) sample for pathological diagnosis. The rates of success in terms of obtaining a sample depended on the randomisation order of the biopsies and whether or not they were done in conjunction with hysteroscopy (*Table 39*). The Tao brush significantly outperformed the Pipelle in high-risk women, with the difference in the percentages being 35 percentage points (95% CI 26 to 45 percentage points). *Figure 7* illustrates this information graphically (ignoring the randomisation order of the biopsies). Considering all women randomised (ITT), the corresponding difference in percentages was 29 percentage points ($p < 0.001$, 95% CI 21 to 37 percentage points, $n = 200$).

If all moderate-risk women with either biopsy successfully completed were considered, then the Pipelle biopsy performed better than the Tao brush biopsy (*Table 40*, *Figure 8*). The difference of 4.9 percentage points was not quite statistically significant ($p = 0.055$). Considering all women randomised (ITT), the corresponding difference in percentages was 4 percentage points ($p = 0.055$, 95% CI 0 to 8 percentage points, $n = 326$).

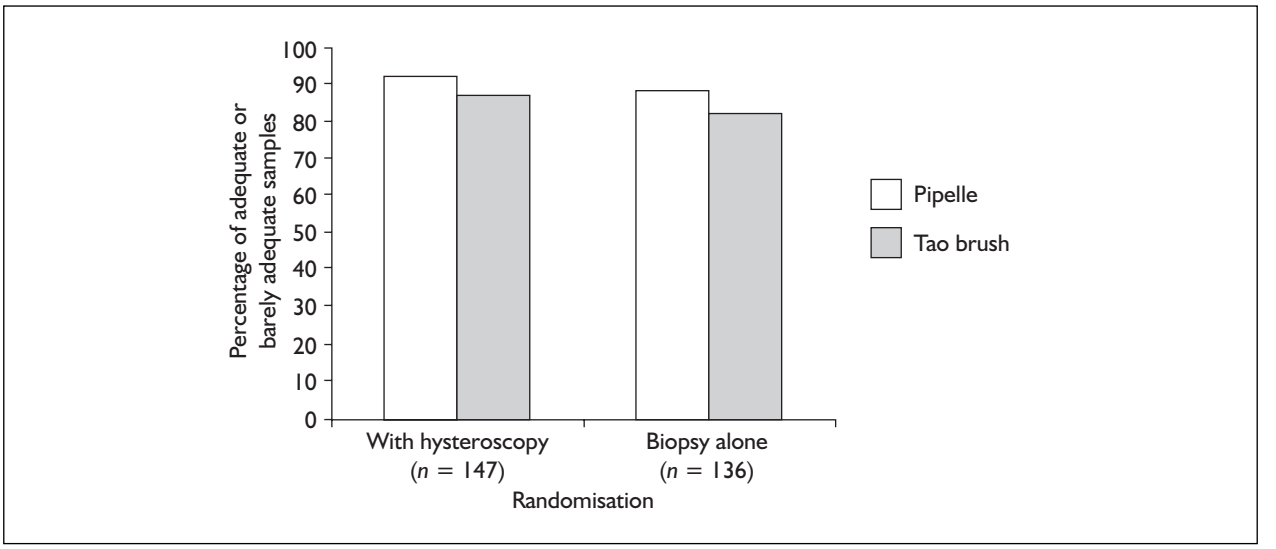


FIGURE 8 *Moderate-risk women: percentage of acceptable samples for Tao brush and Pipelle biopsies: effective* n

Abnormalities detected

Of the ten cases of cancer and one case of CAH detected in this study, all women were randomised to Tao and Pipelle endometrial sampling. For one woman there was failed insertion for both sampling devices, and another woman was the single case of cervical cancer in whom study endometrial biopsies were not undertaken. Of the remaining nine cases (one being complex atypical hyperplasia), seven were detected by both devices and two by Tao brush only (Pipelle providing inadequate samples). Sensitivities and specificities have been presented in *Table 28b*. While Tao was in general shown to provide a significantly higher percentage of adequate samples in postmenopausal women, and detected two cases of cancer missed by Pipelle (out of a total of nine), the small numbers of cancer cases mean that it is not possible to make a conclusive statement regarding diagnostic performance of Pipelle compared with Tao sampler.

Adverse events

The data collection forms did not ask the clinician to attribute any adverse event reported to the Pipelle sampler or Tao brush. Since an adverse event might be expected to occur more often with the first insertion, it was investigated whether adverse events occurred more often for one ordering than the other. However, they were equally distributed among those who had Pipelle or Tao first (for blind biopsy 16 where Pipelle was first and 5 where Tao brush was first, and for biopsy with hysteroscopy 11 and 8, respectively, for Pipelle and Tao first). So there is no suggestion that one or other device was more problematic.

With regard to incomplete biopsies, in the high- and moderate-risk groups, where only one of the two assigned biopsies was completed, probably because of concerns about the patient, there were three women where Tao only was completed and five where Pipelle only was completed, and thus no evidence for a trend.

Patient acceptability

High- and moderate-risk women were randomised to both the Pipelle and the Tao brush biopsies, and therefore a within-patient comparison of relative acceptability can be made. Women were asked their preference (first or second biopsy), but only 277 (63%) women answered the question out of 439 women who had both biopsies successfully completed. It is possible that that some of the non-response is because women either did not realise that two biopsies had been undertaken or did not have a real preference. A total of 91 (33%) women preferred the Tao brush biopsy, compared with 50 (18%) women who preferred the Pipelle biopsy. This difference was statistically significant (McNemar's test, $p < 0.001$). As can be seen in *Figure 9*, a greater proportion of both groups of women preferred the Tao brush biopsy.

Low-risk women were randomised to either the Pipelle or Tao brush biopsy, so their experiences of the biopsy can be compared between the two different groups of women. However, there is low statistical power because of the small numbers. The proportion of women who considered their biopsy to be markedly unpleasant was very similar across the two methods (26% of 23 women for Tao brush and 21% of 29 women for Pipelle). This

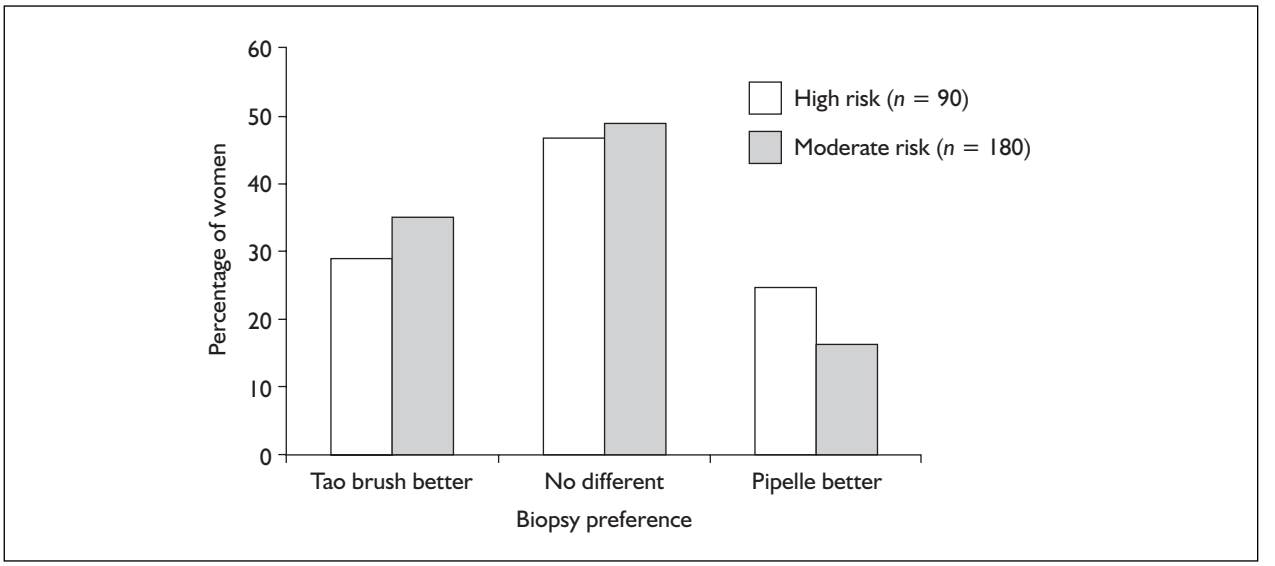


FIGURE 9 *High- and moderate-risk women: preference for biopsy method*

TABLE 41(a) *Resource use to last investigation: appointments, letters and outpatient investigations*

Risk group	Randomised investigation	Mean number of events per woman in subgroup			
		Appointments	Letters	Study outpatient investigations	Extra outpatient investigations
High	H+B (*n* = 100)	1.39	2.45	2.92	1.15
	U+B (*n* = 97)	1.56	2.30	2.80	1.41
Moderate	B (*n* = 78)	1.72	2.18	1.97	2.01
	H+B (*n* = 82)	1.77	2.82	2.91	1.67
	U+B (*n* = 77)	1.91	2.43	2.86	1.57
	U+(H+B) (*n* = 80)	1.63	2.66	3.73	1.54
Low	None (*n* = 59)	1.88	2.49	0.00	1.75
	B (*n* = 32)	1.91	2.34	1.00	1.75
	H+B (*n* = 30)	2.57	3.57	1.80	2.90
	U (*n* = 32)	1.41	1.59	0.97	2.47
All women	(*n* = 667)	1.71	2.47	2.39	1.67

was also true for the proportion of women suffering markedly from cramp or discomfort (45% of 22 women for Tao brush and 52% of 29 women for Pipelle).

Cost analysis

Resource use

Each evaluation method was costed to completion of investigations. This involved extraction of dated event data from case notes for up to 2 years. Each event was coded for type and variant, for example, giving zoladex would be of type 'drug treatment' and variant 'luteinising hormone-releasing hormone (LH-RH) analogue'. Costs were established for each variant. If a particular variant had a cost distinct from other similar variants then it retained a distinct code. The mean number of events of resource use, for the main types, is summarised in *Table 41(a,b)*.

The expected number of study investigations in the high-risk group was three (one visualisation plus two biopsies), in the moderate-risk group two, three or four (the latter being hysteroscopy plus biopsy plus TVUS with two visualisations and two biopsies) and in the low-risk group none, one or two. The actual mean number of investigations per woman differed from the expected because a proportion of women did not undergo all of their study investigations. Extra investigations were more plentiful the lower the risk group.

Inpatient investigations were most common in the high- and low-risk women, and drug treatments

TABLE 41(b) *Resource use to last investigation: inpatient investigations and drug treatments*

Risk group	Randomised investigation	Mean number of 'events' per woman in subgroup	
		Inpatient investigations	Drug treatments
High	H+B (n = 100)	0.37	0.16
	U+B (n = 197)	0.53	0.16
Moderate	B (n = 78)	0.15	0.51
	H+B (n = 82)	0.33	0.46
	U+B (n = 77)	0.26	0.57
	U+(H+B) (n = 80)	0.15	0.53
Low	None (n = 59)	0.53	0.63
	B (n = 32)	0.25	0.56
	H+B (n = 30)	0.37	0.80
	U (n = 32)	0.28	0.41
All women	(n = 667)	0.33	0.43

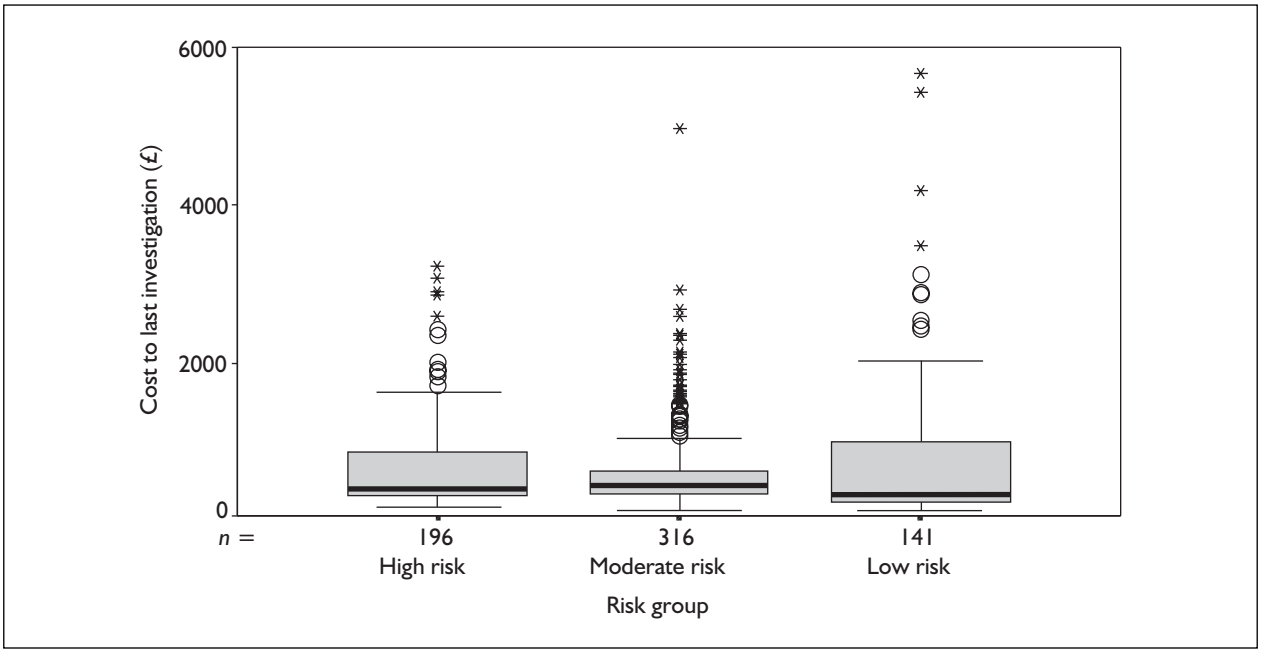


FIGURE 10 *Costs to last investigation by risk group*

were least common in the high-risk women. The greatest impact on cost arises from inpatient investigations, then outpatient investigations and clinic appointments. *Figure 10* shows the distribution of costs by risk group.

Comparison of cost-effectiveness of the three methods of endometrial evaluation

Costs were examined separately within risk groups. This is mainly because the significance of bleeding symptoms varies by risk group, but also because the nature and impact of the bleeding symptoms differ across the groups. In the high-risk group the main concern is to exclude cancer, whereas in the two lower risk groups there is (also or mainly) a need to alleviate the bleeding symptoms. As shown in *Figure 10*, the group with lowest intensity of assigned investigations, the low-risk group, while having the lowest median cost, also had some of the highest costs to last investigation.

Table 42 presents the mean cost-effectiveness to last investigation for the high- and moderate-risk groups, with sensitivity limits. For each mean three sensitivity analyses were undertaken, separately varying the unit costs of hysteroscopy, Pipelle sampler and Tao brush by plus or minus 10%. (This was not done for ultrasound as its cost is well established.) There were significant differences in cost-effectiveness for the investigation options within each risk group (Kruskal–Wallis test, $p < 0.001$ for high risk, $p < 0.0001$ for moderate risk). In the high-risk group hysteroscopy (with

TABLE 42 *Moderate- and high-risk women: cost-effectiveness (£) to end of investigation by randomised options (including all treatments)*

Randomised investigation	High risk		Moderate risk	
	n	Mean cost (sensitivity limits)	*n*	Mean cost (sensitivity limits)
B			78	£474 (458–493)
H + B	99	£632 (602–662)	81	£686 (656–716)
U + B	97	£720 (700–740)	77	£602 (583–621)
U + H + B			80	£638 (607–668)
p (Kruskal–Wallis test)		0.003		< 0.0001

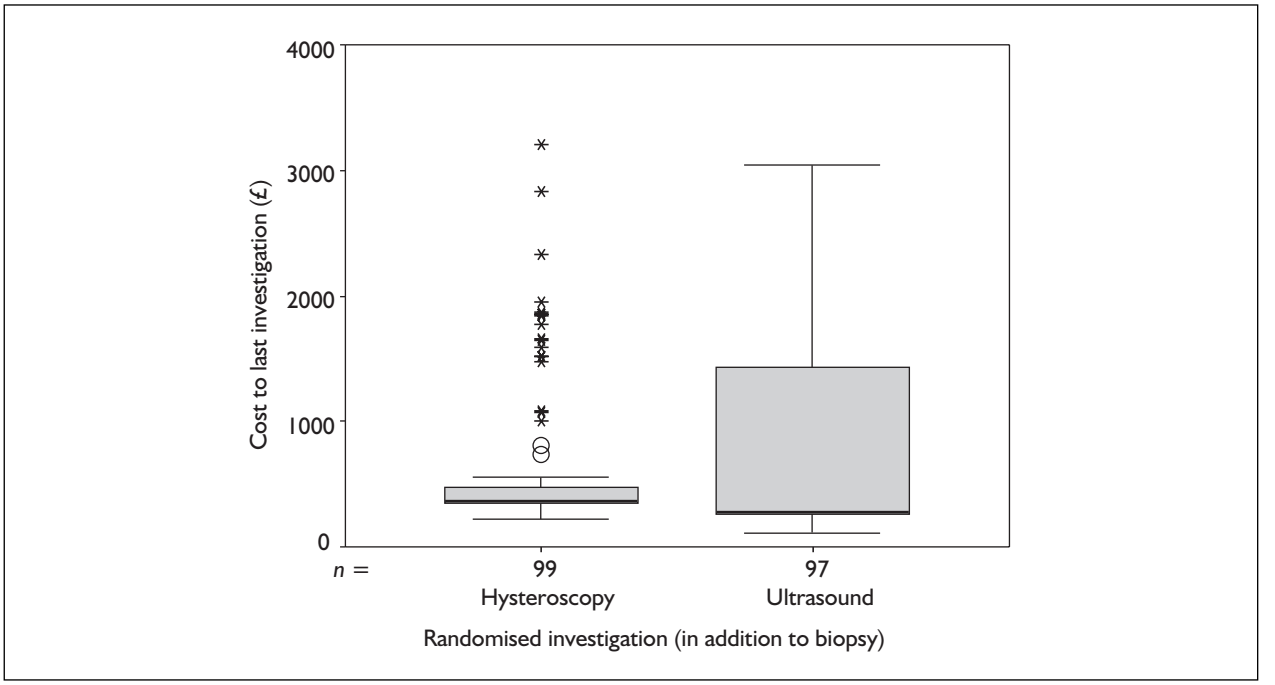


FIGURE 11 *High-risk women: cost to last investigation by randomised investigation*

biopsy) was more cost-effective than ultrasound and biopsy. In the moderate-risk group the significant difference arose mainly because biopsy on its own was so much more cost-effective than the visualisation options separately or in combination. The distributions of individual patient costs for these two groups are shown in *Figures 11 and 12*.

The number in the low-risk group are so small that the cost-effectiveness was summarised with Tao and Pipelle data combined as simply biopsy (unspecified). *Table 43* presents the cost-effectiveness to last investigation for the low-risk group.

There was a significant difference in cost-effectiveness among the investigation options ($p < 0.0001$). Two of the options, ultrasound and blind biopsy on its own, were considerably more cost-effective than hysteroscopy (with biopsy) or initially no investigation. The distribution of cost-effectiveness is illustrated in *Figure 13*. The difference in unit cost for Tao compared with Pipelle is an extra £29.

TABLE 43 *Low-risk women: cost (£) to end of investigation by randomisation option*

Randomised investigation	*n*	Cost (sensitivity limits)
None	59	£748 (739–755)
P/T	32	£452 (443–461)
H + P/T	30	£830 (801–855)
U	32	£378 (376–381)
p (Kruskal–Wallis test)		<0.0001

To explore the reason for the higher costs in some subgroups than in others, the cost to last investigation excluding treatments was calculated. This is shown in *Figure 14*. In addition, the investigation costs assigned by the study

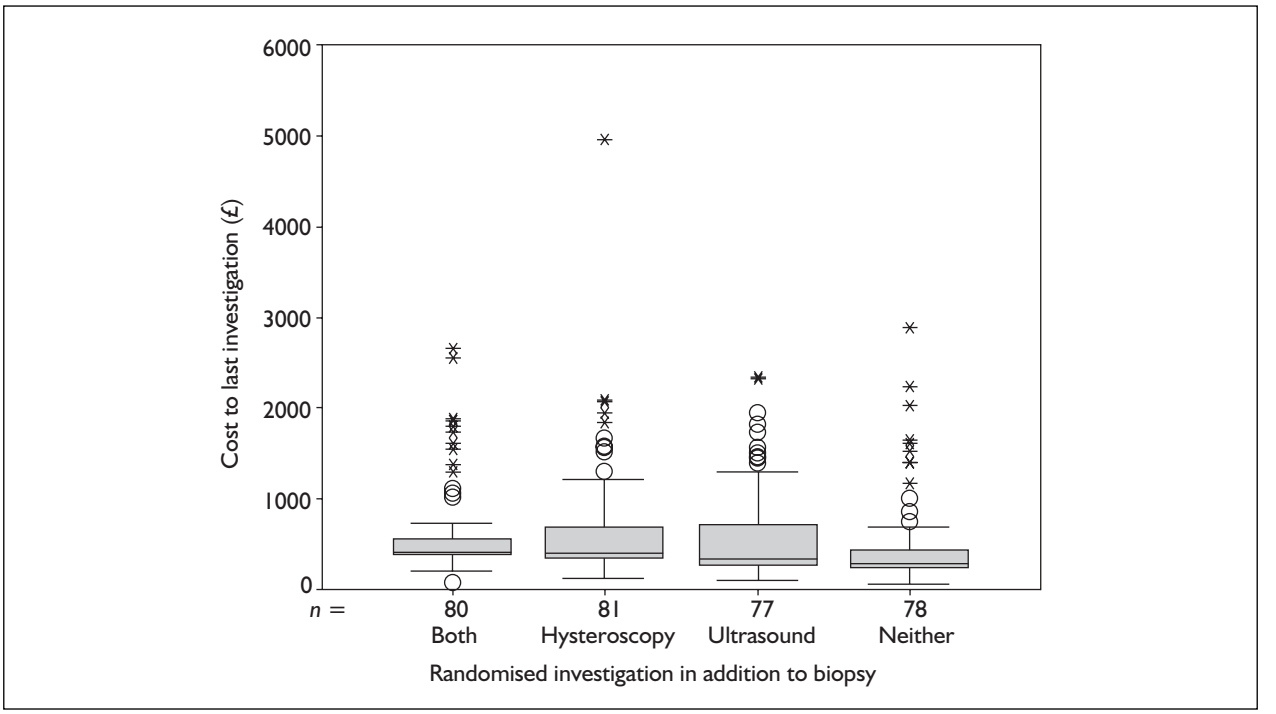


FIGURE 12 *Moderate-risk women: cost to last investigation by randomised investigation*

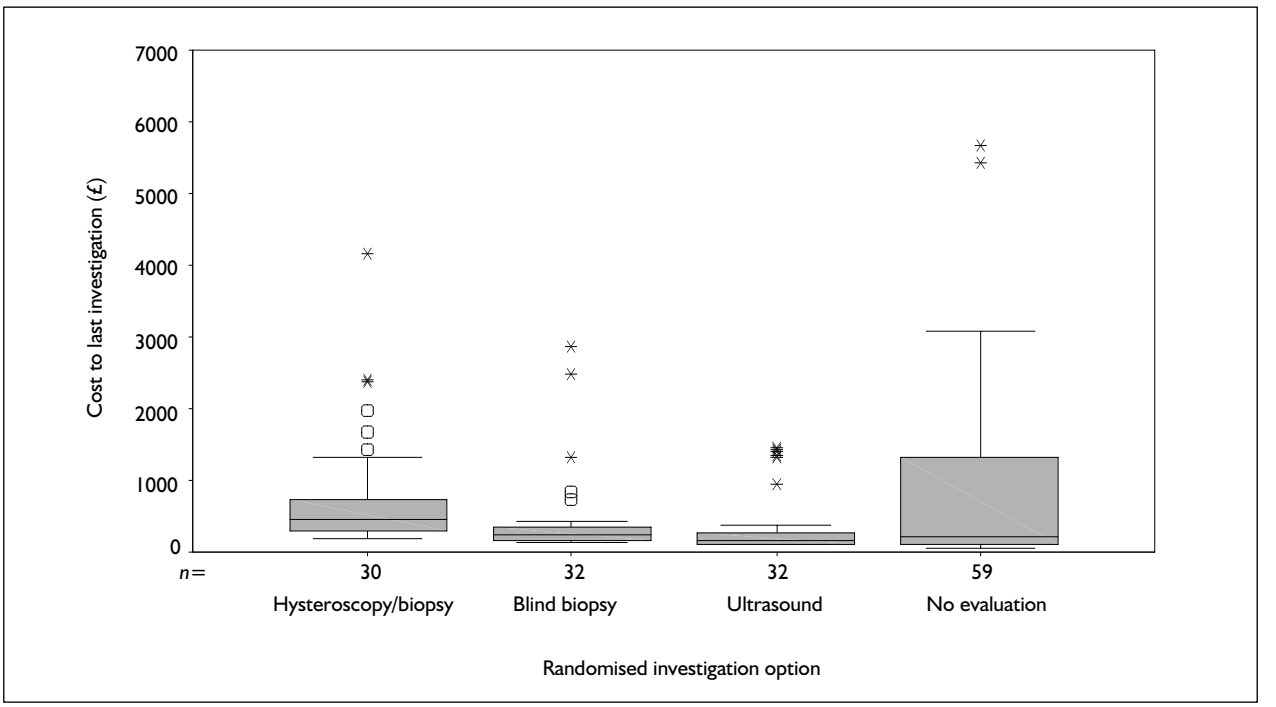


FIGURE 13 *Low-risk women: cost-effectiveness by randomisation option*

were partitioned out, separately from additional investigation costs. Treatment costs made a marked difference to total cost in some subgroups, especially in the low-risk group, but excluding these did not notably alter the relative cost-effectiveness of investigations within groups.

Cost-effectiveness for Pipelle sampler compared with Tao brush sampler

The design of the study, with paired comparison of the Tao brush and Pipelle endometrial sampler, did not allow a straightforward comparison of the cost-effectiveness of the two devices, since if the Pipelle failed to obtain an adequate sample, as it

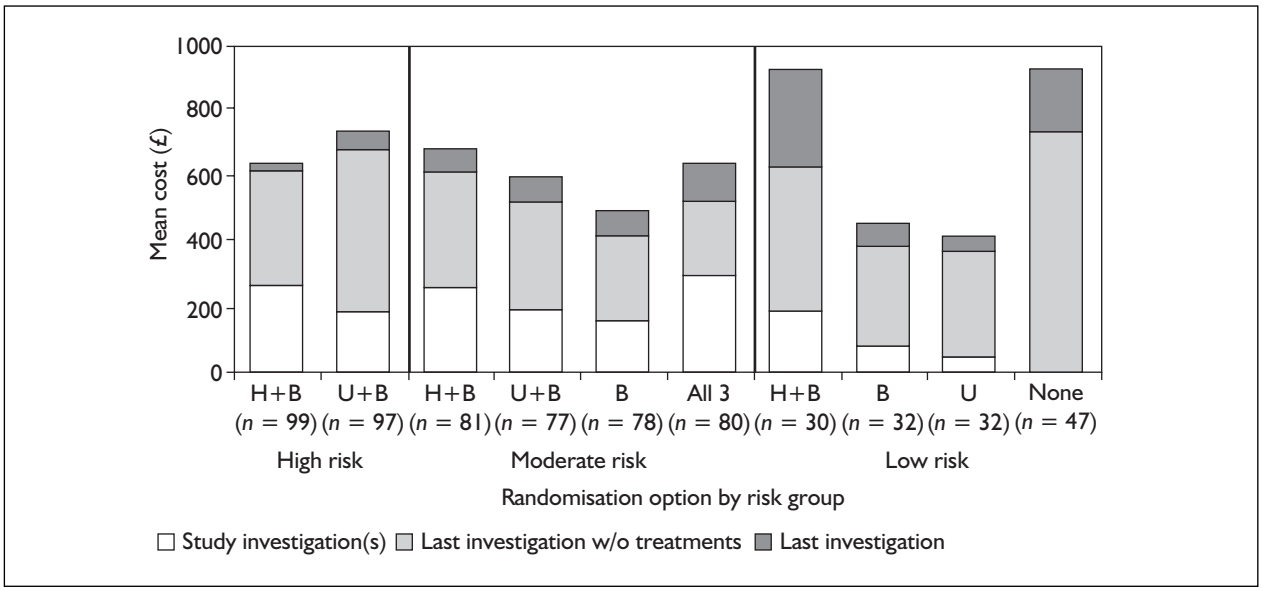


FIGURE 14 *Cost to last investigation including and excluding treatments, by risk group and randomised investigation*

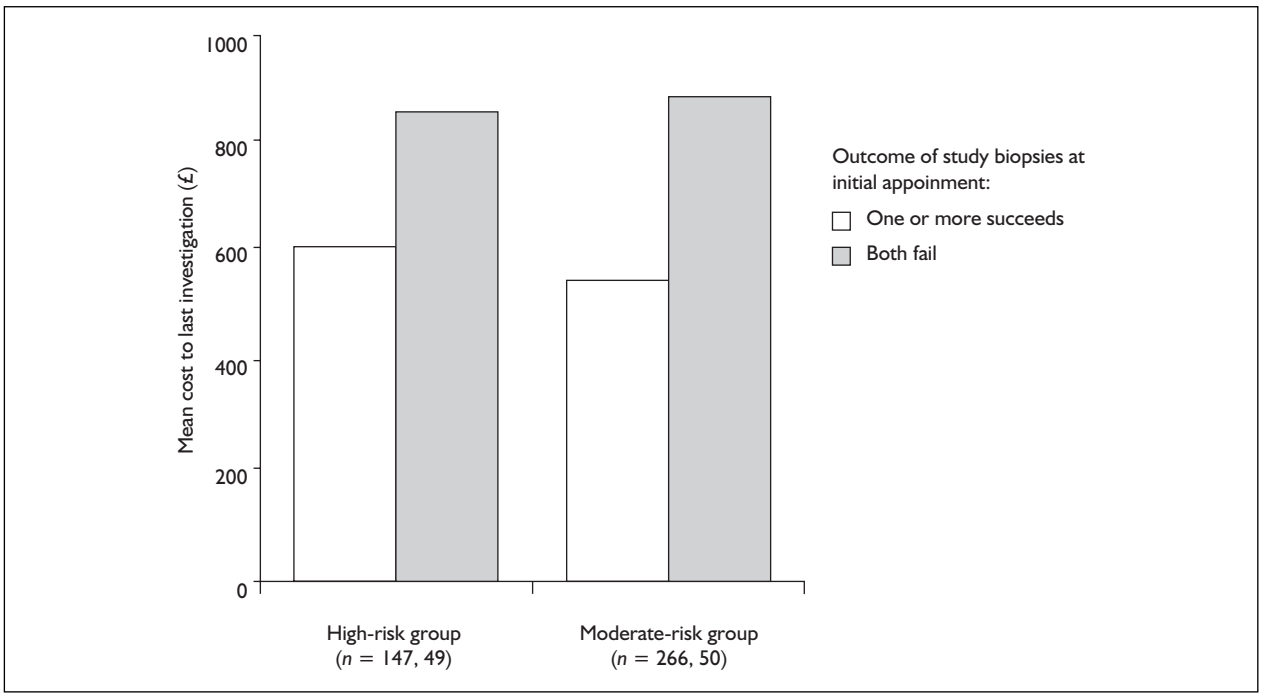


FIGURE 15 *Mean cost to last investigation by risk group, for women where at least one study biopsy obtained an adequate sample, compared with those where both study biopsies failed*

often did in the high-risk group, then if the Tao brush had provided a sample, there would have been no need for the clinician to take any further action (and incur costs) to make good the lack of an histological diagnosis. As explained in Chapter 3, the case-note review data were used to estimate in the high- and moderate-risk groups the additional costs that would have been likely to be incurred if Pipelle only had been used for first line biopsy sampling. *Figure 15* shows that in this study the costs for women for whom a first line biopsy is successful were considerably less than costs for women where both biopsies failed. The additional costs accruing because of failed initial biopsy were if anything higher in the moderate-risk women, despite the lower risk of endometrial cancer in the premenopausal women.

These calculations provide estimates of the costs that might apply if Pipelle alone were the first line

endometrial sampler, for the two scenarios where the first line Pipelle succeeded in obtaining an adequate sample, and the situation where it failed. This is estimated on the basis that no other outpatient sampling device is feasible or available, so that in cases where concern is high some inpatient investigation may be required to obtain a sample. In the high-risk group the Pipelle sampler often failed when a Tao brush would have been successful in terms of sample adequacy (63 women out of 200 investigated, 32%). Therefore, where Pipelle is the first line sampler but fails to obtain an adequate sample despite successful insertion (48% of high-risk women), a potentially cost-effective strategy would be to try the Tao as next line biopsy method. There would be an expectation of success (and cost savings) in over three-quarters of women with failed Pipelle (high-risk women). Cost savings would be more secure, and there would be less inconvenience and worry for the woman, if the lack of adequate Pipelle sample could be judged by visual inspection by the operator, and a fallback Tao sample taken immediately. The cost for this intermediate scenario could be calculated by adding to the 'if biopsy succeeds' costs shown in *Figure 15*, the marginal additional cost of a Tao investigation at the same appointment (approximately £100). Further analysis of the case-note review data will allow modelling of the costs for a range of scenarios using a single method of biopsy.

Secondary objectives

Factors influencing incomplete, unsuccessful and unpleasant investigations

Incomplete investigations

Parity was an important factor determining whether or not the Pipelle or Tao brush biopsies

TABLE 44 *High- and moderate-risk women: factors associated with failed insertions for the Pipelle and Tao brush biopsies: effective* n

Factor	Level	Pipelle biopsy		Tao brush biopsy	
		Failed insertion/women (%)	*p* (χ^2 test)[a]	Failed insertion/women (%)	*p* (χ^2 test)[a]
Parity	Nulliparous	15/68 (22.1)	0.001	14/69 (20.3)	0.005
	Parous	34/420 (8.1)		35/418 (8.4)	
HRT use	No	59/404 (14.6)	0.325	59/402 (14.7)	0.317
	Yes	8/82 (9.8)		8/82 (9.8)	
Menopausal status (risk group)	High	31/195 (15.9)	0.296	31/195 (15.9)	0.309
	Moderate	39/319 (12.2)		39/317 (12.3)	
Age group (years)	23–44	19/136 (14.0)	0.317[a]	20/135 (14.8)	0.451[a]
	45–49	13/144 (9.0)		13/144 (9.0)	
	50–54	20/123 (16.3)		19/122 (15.6)	
	55–86	18/111 (16.2)		18/111 (16.2)	

[a] χ^2 test for trend.

TABLE 45 *High-risk women: factors associated with inadequate samples separately for the Pipelle and Tao brush biopsies: effective* n

Factor	Level	Pipelle biopsy		Tao brush biopsy	
		Inadequate sample/ women (%)	*p* (χ^2 test)	Inadequate sample/ women (%)	*p* (χ^2 test)
Parity	Nulliparous	11/21 (52.4)	0.81	3/21 (14.3)	0.78
	Parous	66/141 (46.8)		17/141 (12.0)	
HRT use	No	55/108 (50.9)	0.16	11/108 (10.2)	0.23
	Yes	20/53 (37.7)		9/53 (17.0)	
Age group (years)	23–49	6/18 (33.3)	0.01[a]	1/18 (5.6)	0.03[a]
	50–54	20/56 (35.7)		3/56 (5.4)	
	55–86	52/90 (57.8)		16/90 (17.8)	

[a] χ^2 test for trend.

TABLE 46 *Moderate-risk women: factors associated with inadequate samples separately for the Pipelle and Tao brush biopsies: effective* n

Factor	Level	Pipelle biopsy		Tao brush biopsy	
		Inadequate sample/ women (%)	*p* (χ^2 test)	Inadequate sample/ women (%)	*p* (χ^2 test)
Parity	Nulliparous	8/32 (25.0)	0.001	11/34 (32.4)	0.001
	Parous	13/245 (5.3)		24/242 (9.9)	
HRT use	No	20/237 (8.4)	0.99	30/235 (12.8)	0.93
	Yes	2/21 (9.5)		2/21 (9.5)	
Age group (years)	23–44	8/115 (7.0)	0.20[a]	14/113 (12.4)	0.99[a]
	45–49	8/115 (7.0)		15/115 (13.0)	
	50–86	7/50 (14.0)		6/50 (12.0)	

[a] χ^2 test for trend.

failed in terms of insertion in high- and moderate-risk women. HRT use, menopausal status and age did not have an influence on this aspect of failure (*Table 44*).

The rates of success in terms of obtaining a sample depended on the randomisation order of the biopsies and whether or not they were done in conjunction with hysteroscopy (*Table 39*).

Inadequate samples

Tables 45 and 46 show patient factors associated with inadequate sample collection in high- and moderate-risk women. Therefore, in high-risk women where insertion was possible, inadequate samples with Pipelle were more likely if the woman was over 55 years of age, and there was a trend for their being more likely if there was no HRT use. For the Tao brush, failure to obtain an adequate sample was fairly low overall, but slightly more likely if the woman was over 55 years of age.

For the moderate-risk women for whom insertion had been possible, the only patient factor associated with adequate sample was parity, and this was important for both Tao and Pipelle samplers. In high- and moderate-risk women for whom one or both biopsies were completed, paired analysis of discordance in outcome with Tao and Pipelle samplers showed that the strongest predictor of which sampler would have inadequate samples was menopausal status ($p = 0.002$). This was in the direction of the Tao brush significantly outperforming the Pipelle sampler in postmenopausal women. After adjusting for menopausal status, there was only a borderline effect of parity on discordance ($p = 0.11$, Tao performing slightly better than Pipelle in nulliparous subgroups of both risk groups), and no further effect due to age group ($p = 0.18$), or HRT use ($p = 0.76$).

Predicting which women suffer from discomfort or bleeding after investigation

Stepwise logistic regressive revealed that those who were more likely to suffer from abdominal discomfort following their clinic visit were younger women (OR 0.58 for each successive increase in age by 10 years, 95% CI 0.46 to 0.73), women randomised to hysteroscopy (OR 1.87, 95% CI 1.25 to 2.80), women randomised to biopsy (OR 5.99, 95% CI 2.91 to 12.3), women with higher GHQ A somatic scores (OR 2.05 for increase in scale score by 10 points, 95% CI 1.23 to 3.42) and women in poorer health (OR 1.55 for each successive increase by one point of the four-point scale, 95% CI 1.10 to 2.19).

Analyses of bleeding after the clinic visit showed that age, being randomised to hysteroscopy and current health were not significant predictors of bleeding. However, being randomised to ultrasound and the woman's risk group were predictors. The most important risk factor was being randomised to biopsy (OR 6.62, 95% CI 2.62 to 16.8), followed by higher GHQ A somatic scores (OR 2.44 for a ten-point increase in scale score, 95% CI 1.49 to 3.97), being randomised to ultrasound (OR 0.59, 95% CI 0.39 to 0.88), low-risk women compared with high-risk women (OR 2.25, 95% CI 0.99 to 5.12) and moderate-risk women compared with low-risk women (OR 1.35, 95% CI 0.60 to 3.05). The fact that being randomised to ultrasound was found to be protective may have arisen because many of the women who were randomised to ultrasound were not randomised to hysteroscopy, and so this result could reflect an increased risk of finding the investigation unpleasant if randomised to hysteroscopy.

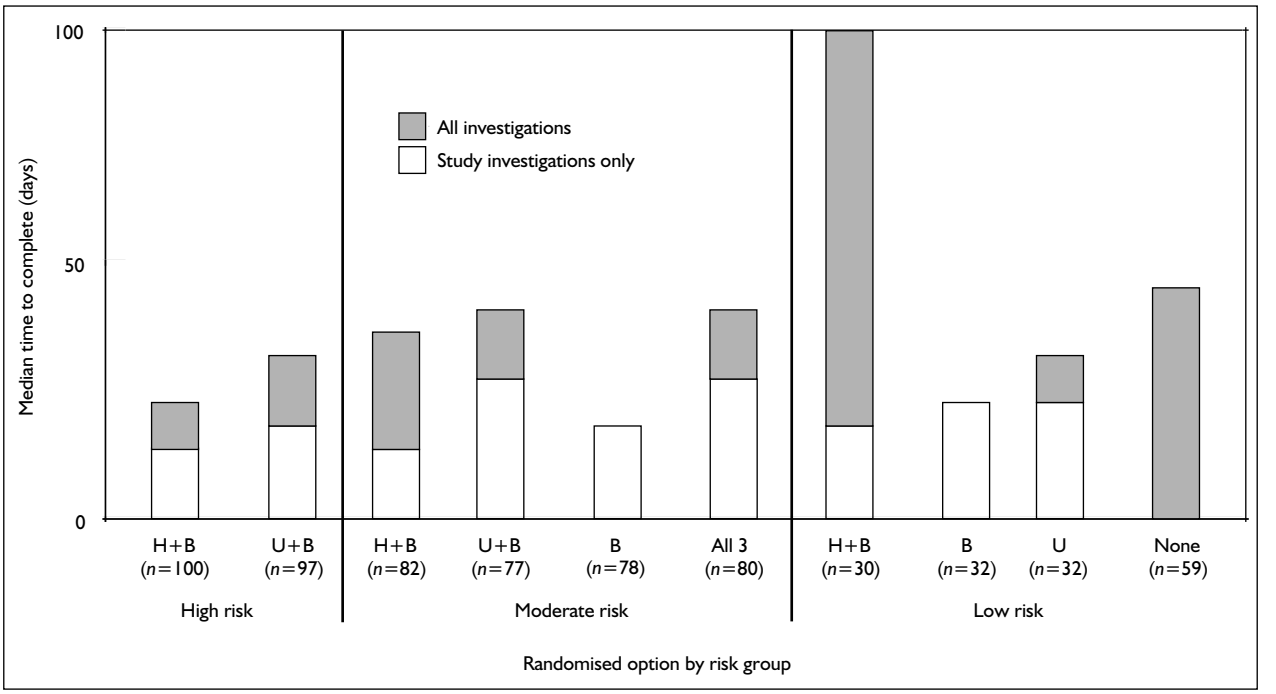


FIGURE 16 *All women: median time taken to complete study investigations, and all investigations, by randomisation and risk group*

Pattern of adjunctive evaluations undertaken and resource use across time

Some understanding of the factors contributing to costs can be obtained by examining the time to last investigation. *Figure 16* shows by randomised investigation within risk groups the median number of days to complete all investigations, and within that the median number of days to complete the randomly assigned study investigations.

Where the randomisation was to biopsy only this could be undertaken at the first clinic appointment, so no time was required for completion of study investigations. The time taken for hysteroscopy and ultrasound study investigations was dependent on the waiting lists, and if more sessions could have been made available these could have been shorter. In all randomisation arms the median time to complete all investigations was greater than the time to complete study investigations. The low-risk subgroups comprise small numbers of women, so the median times to complete all investigations for these subgroups should be interpreted with caution. The data on resource use across the timespan of the study are still to be analysed.

Clinician attitudes to investigations over the course of the study

Clinician assessments were developed by undertaking interviews with a selection of gynaecologists at the beginning of the study, and informed by these interviews and findings from the literature, developing a succinct questionnaire assessment of evaluation strategy preference. This was sent to all consultants and senior registrars at the RIE around month 23 of the study (after exposure to the use of all three study evaluations in their patients). All clinicians were surveyed again at month 40 (after the preliminary study report on the findings for evaluation efficiency had been circulated to all clinicians at the RIE). Selected findings are presented in Appendix 7.

Ancillary analyses

Psychological and personality factors related to worry about health and preferences regarding consultation style

Table 47 presents correlations of the patient characteristics shown in *Table 17* with the NEO personality factors and the GHQ total score and A scale score (somatic symptoms). The intercorrelations between the NEO factors and GHQ scales are also shown.

The dimension of personality labelled neuroticism (N), the total GHQ score and the GHQ somatic symptoms subscale (A) score were most strongly associated with worrying about one's health, but also with finding bleeding worrying, thinking something seriously wrong, judging one's health worse than others and sensitivity to pain. The GHQ somatic symptoms (A) subscale only was associated with

TABLE 47 *Non-parametric correlations between patient characteristics and NEO and GHQ scores*

	NEO					GHQ	
Self-report measures	**N**	**E**	**O**	**A**	**C**	**Total**	**A**
Strength of agreement that … doctor should decide treatment			–0.25**		0.10*		
she likes to be given choice about tests/treatment			0.19**			0.10*	
she likes to be told as much as possible about condition		0.16**	0.22**	0.13**	0.21**	0.10*	0.10*
bleeding problems are very worrying	0.08*		–0.19**			0.13*	0.10*
Degree of worry about health	0.34**	–0.13**		–0.12*	–0.18**	0.38**	0.35**
How much think something seriously wrong with body	0.21**		–0.16**	–0.11*	–0.10*	0.30**	0.31**
How 'likely' that bleeding may be due to cancer					–0.12*		0.10*
Judging own general health worse than other women of same age (rather than same or better)	0.24**	–0.16**	0.10*		–0.21**	0.32**	0.34**
Sensitivity to pain	0.21**	–0.13*	–0.08*		–0.13**	0.17**	0.14**
NEO scores							
Neuroticism	~	–0.41**	–0.12*	–0.27**	–0.39**	~	~
Extraversion	~	~	0.18**	0.19**	0.40**	~	~
Openness	~	~	~			~	~
Agreeableness	~	~	~	~	0.28**	~	~
Conscientiousness	~	~	~	~	~	~	~
GHQ scores							
A: somatic symptoms	0.38**	–0.14**		–0.14**	–0.18**	~	~
B: anxiety	0.56**	–0.15**		–0.17**	–0.22**	0.88**	0.60**
C: social dysfunction	0.37**	0.21**	–0.10*		0.23**	0.66**	0.48**
D: depression	0.65**	–0.29**		–0.14**	–0.27**	0.64**	0.37**
Total	0.59**	–0.22**		–0.18**	–0.26**	~	0.83**

Spearman correlations: *$p < 0.05$, **$p \leq 0.001$, empty cells denote $p > 0.05$, ~ denotes correlation is 1.0 or shown elsewhere in the table.
640, 619 and 578 women completed the Health, NEO and GHQ questionnaires respectively. The cancer item had 26% missing responses but for all other items missing responses were fewer that 4%.

judgements that the bleeding was likely to be due to cancer. In contrast to the neuroticism dimension on the NEO, the conscientiousness (C), extraversion (E) and agreeableness (A) personality dimensions were associated with denial of worry about health, and on the whole with judgements of better health than others and denial of sensitivity to pain, thinking something seriously wrong and thinking that the bleeding may be cancer. Women with the most open (O) personalities tended to judge their health worse than others, but they did not tend to find bleeding problems worrying or tend to think that there was anything seriously wrong.

Thinking that bleeding problems are worrying, or that bleeding may be due to cancer, showed different patterns of correlation to more general worries about health, although still, albeit less strongly, related to GHQ total and somatic symptoms scores. Conscientious women were least able to acknowledge the possibility that bleeding meant cancer, and open women least likely to agree that bleeding problems were very worrying. Agreeing that bleeding was worrying showed a different response pattern depending on whether the questionnaire was completed at the clinic, or later at home. If the questionnaire was completed at the clinical then there was the same pattern of correlations, but stronger, whereas if it was completed at home then there was association only with agreeableness and with the depression scale of GHQ.

Wanting to be fully informed about the condition was associated with all personality dimensions other than neuroticism (N), that is with openness (O), conscientiousness (C), extraversion (E) and agreeableness (A), and also with GHQ total and somatic (A) subscale scores. Higher openness scores and higher GHQ total scores were associated with reporting a wish to be given choice as to tests and treatments.

While women with high scores on agreeableness or conscientiousness tended to agree that the doctor should decide treatment, women with the most open personalities disagreed strongly. If the questionnaire was completed at home, open-minded women disagreed even more strongly (rho = –0.295), and women with high scores for agreeableness or conscientiousness did not tend to agree that the doctor should decide.

Further analyses will be undertaken to examine the relevance of individual factors regarding bleeding symptoms, ease of reassurance and satisfaction with care.

Summary of results

Recruitment and participant flow

- Recruitment was adequate in the high- and moderate-risk groups, but poor in the low-risk group, the latter due to a shortage of referrals, a reluctance among the women to participate and a reluctance among the managing clinicians to give consent to randomisation. These recruitment factors and the resulting small subgroup numbers make it difficult to interpret the findings in the low-risk group. In the moderate- and high-risk groups 68% of those approached about the study participated.
- Over 90% of women completed their recruitment questionnaires and report on clinic visit, 82% completed all of their randomised investigations and over 83% returned their review of the clinic visit.
- There were high rates of follow-up and case-note review to 10 months, and given the fact that some participants had not, by the end of the data collection, been enrolled for 2 years, adequate follow-up and review to 2 years.

Comparison of the three methods of endometrial evaluation

Visualisation

- In the high-risk group there was a non-significant trend for successful visualisation being more frequent with ultrasound than hysteroscopy (87% versus 79%), although this was mainly through failures to insert the hysteroscope.
- The situation was similar but with a significant advantage for ultrasound in both the moderate-risk group (89% versus 77%, $p < 0.005$) and the low-risk group (97% versus 65%, $p = 0.003$).
- When the data for high- and moderate-risk women were combined there were significant differences overall for visualisation success in favour of ultrasound ($p = 0.002$).

Adequate sample

- Failed insertion of biopsy device was equally common for the Pipelle sampler and Tao brush (13% for both) in high- and moderate-risk women, and for both devices significantly more common for nulliparous than for parous women ($p = 0.001$ for Pipelle and $p = 0.005$ for Tao).
- Adequate biopsy samples were obtained in 75–78% of cases in the moderate- and low-risk women (ITT, summing across randomisation options within group), with very little difference by group or biopsy method, but in high-risk women adequate samples were much more frequent when using a Tao brush rather than a Pipelle sampler (72% versus 43%, $p < 0.0001$).
- In the high-risk group there was an improvement in rates of adequate samples if the biopsy had been undertaken at the same time as hysteroscopy (non-blind), and this was most striking for Tao brush (83% versus 61%, $p < 0.001$). There was a similar trend in the moderate-risk group, but it was not significant for either the Pipelle sampler or Tao brush. However, when the data for high- and moderate-risk women were combined there were significant differences overall for adequate samples in favour of their being taken at the same time as hysteroscopy, for both Pipelle ($p = 0.027$) and Tao brush ($p = 0.002$).
- Where insertion was possible, inadequate samples in high-risk women were more common in the oldest women by Pipelle, and those in moderate-risk women were most common in nulliparous women for Pipelle and Tao.
- In high- and moderate-risk women for whom one or both biopsies were completed, paired analysis of discordance in outcome with Tao and Pipelle samplers showed that the strongest predictor of which sampler would have inadequate samples was menopausal status ($p = 0.002$). This was in the direction of the Tao brush significantly outperforming the Pipelle sampler in postmenopausal women.

Abnormalities detected and adverse events

- For all risk groups combined, more polyps were detected with hysteroscopy (12.5% versus 4.4%, $p < 0.001$) and more fibroids by ultrasound (32% versus 13.1%, $p < 0.001$). This pattern was stronger in the high-risk group (17% versus 4% for polyps, $p = 0.006$ and 29% versus 7% for fibroids, $p < 0.001$).
- In total, nine cases of endometrial cancer, one case of CAH of the endometrium and one case of cancer of the cervix were detected. The rate of endometrial disease detected (cancer or CAH) was 3.0% in the high-risk (postmenopausal) group and 1.2% in the moderate-risk (premenopausal) group.
- Of the six cases of endometrial disease in the high-risk group, only three were randomised to TVUS and two of these were detected as possible cancer owing to endometrial thickness greater than 4 mm (sensitivity 67% CI 21 to 94%, $n = 95$). The woman with cancer detected had a measured endometrial thickness of 1.9 mm.
- All adverse events were minor and there were none with ultrasound. Overall there was a significant difference in rates across the three methods, with 16% of adverse events occurring with hysteroscopy and 10% with biopsy.

Acceptability

- Women reporting abdominal discomfort after investigation were more likely to be young, to have been investigated by hysteroscopy or biopsy, to have high scores on the GHQ somatic scale and to have self-reported poorer health.
- Women reporting bleeding afterwards were more likely to be premenopausal, to have been randomised to biopsy, to have high scores for GHQ somatic scale and not to have been randomised to ultrasound (protective factor).
- In all risk groups and overall, hysteroscopy and biopsy were more likely than ultrasound to be rated as markedly unpleasant (27%, 29% and 11%, respectively).
- On the review of clinic visit questionnaire, completed once study investigations had been completed, the majority of women stated their clinic visit to have been very reassuring (88%, 87% and 83% in the high-, moderate- and low-risk groups, respectively).
- In contrast, lower percentages of women judged that the clinic visit had been very or extremely worthwhile. This was not related to investigation but was related to risk group, with rates decreasing with risk: 77% (ITT) in the high-risk group, 63% in the moderate-risk group and 46% in the low-risk group.
- Before their clinic appointment 65% of women strongly agreed with the statement 'I like to be told as much as possible about my condition', but afterwards 14% of women responded that they did not understand what the doctor told them about their bleeding.
- Women who had hysteroscopy tended to be less discontented with the extent of investigation, and those who had blind biopsy most discontented (32% of moderate- and low-risk women randomised to biopsy only would have liked more investigation).

Outcome at 10 and 24 months

- At the 10-month follow-up for the high-risk group, women investigated by hysteroscopy had better outcomes than women given ultrasound. However, this trend had all but disappeared by 24 months.
- In the moderate-risk group, using a factorial comparison, there appeared to be an advantage separately for hysteroscopy and for ultrasound, but this arose mainly because those randomised to biopsy only (who had had neither visualisation) had the most negative judgements of their clinic experience and outcome, and in particular would have preferred more investigation (42% at 10 months and 38% at 2 years).
- Compared with postmenopausal women, the outcome for the moderate-risk women (with menstrual bleeding problems) was strikingly worse: over half at 10 months and over one-third at 24 months reported that symptoms had not improved, and at both follow-ups about one-quarter believed that their problem had not been cured.

Comparison of two methods of biopsy: Tao brush versus Pipelle sampler

- In high-risk women adequate samples were far more likely to have been obtained by Tao brush than by Pipelle sampler (88% versus 52%, $p < 0.001$ for those where a sample was taken, 72% versus 43% for ITT), but in moderate-risk women this was reversed, with Pipelle marginally better than Tao (91% versus 86%, $p = 0.055$ for those where a sample was possible, 79% versus 75% for ITT). In both groups the differences were slightly more marked if the biopsy was undertaken non-blind, that is, with hysteroscopy.
- Of the ten cases of endometrial disease (one CAH and the rest cancer), all of which were randomised to both biopsy methods, nine were detected by Tao brush, the remaining case being missed because of failed insertion (sensitivity

90% CI 60 to 98%, $n = 523$). For the Pipelle sampler the same case was missed owing to failed insertion, two further cases were missed because of inadequate samples, and the remaining seven cancers were successfully detected (sensitivity 70% CI 40 to 89%, $n = 528$).
- Women who received the two biopsies were asked to say which they preferred (simply as the first or second sample taken), and significantly more preferred the Tao brush ($p<0.001$).

Cost-effectiveness

- The most cost-effective form of evaluation for:
  - the high-risk group was hysteroscopy and biopsy at £632 per person, with the least cost-effective method being ultrasound and biopsy at £720 per person ($p<0.001$). The incremental cost of ultrasound over hysteroscopy and biopsy is therefore £88 per person
  - the medium-risk group was blind biopsy at £474 per person, with the least cost-effective method being hysteroscopy and biopsy at £686 per person ($p < 0.0001$ comparing all four arms). The incremental cost of hysteroscopy and biopsy over both biopsies is therefore £212 per person
  - the low-risk group was ultrasound at £378 per person, with the least cost-effective method being hysteroscopy and biopsy at £830 per person ($p < 0.0001$ comparing all four arms). The incremental cost of hysteroscopy and biopsy over ultrasound is therefore £452 per person
- Univariate sensitivity analysis showed that these results are robust to 10% variation in the costing of the hysteroscopy, Pipelle biopsy and Tao biopsy.
- The study design did not allow empirical comparison of the cost-effectiveness of Tao brush and Pipelle endometrial samplers. The difference in unit cost if a Tao brush biopsy was substituted for Pipelle would be an increase of £29 per patient, whereas the additional cost of adding a Tao biopsy to a routine clinic appointment with Pipelle biopsy would be £100. Alternatively, estimations calculated from study data, of additional investigation and healthcare costs accruing in women for whom neither study biopsy was successful, suggested that a strategy of undertaking a Tao brush biopsy when the routine Pipelle biopsy was judged by visual inspection to be inadequate, would reduce overall costs.

Clinician attitudes

- A high proportion of clinicians expressed a lack of experience in use of outpatient hysteroscopy or the Tao brush, but most were comfortable with the use of a Pipelle sampler.
- There was a change over the time-frame of the study in the expected completion rates for Tao, Pipelle, ultrasound and hysteroscopy, in the direction of findings reported for the randomised trial.
- There was also a change over the time-frame of the study in the anticipated success rates for Tao, Pipelle, ultrasound and hysteroscopy investigations, again in the direction of the trial findings.
- By the second survey, clinicians' reported expectations in terms of general completion rates for the investigations were on the whole correct for ultrasound, correct about half the time for hysteroscopy, but underestimated the situation for the Pipelle sampler and markedly underestimated it for the Tao brush.
- By the second survey, clinicians' reported expectations in terms of general success rates for the investigations were on the whole correct for ultrasound, overestimated visualisation rates for hysteroscopy, but markedly underestimated the rate of adequate samples for the Tao brush. In the case of Pipelle, the differing rates of adequate samples obtained in postmenopausal and moderate-risk women (52% and 91%) make comparison with clinician expectations difficult. However, considering the group for which there is the strongest indication for biopsy if there is AUB, the postmenopausal women, the clinician expectations of rates of adequate samples for Pipelle were optimistic.
- For both the Pipelle sampler and the Tao brush, clinicians showed a strikingly broad range of expectations of performance in terms of completion of tests and adequacy of samples obtained, and in general had more optimistic expectations of the Pipelle sampler than of the Tao brush.
- Neither involvement of patients in the study nor the interim study report was considered to have been influential in changing practice.
- Of clinicians completing the second survey, ten (95%) said they would like a clinical pathway for investigation of AUB.
- Three-quarters of clinicians indicated an interest in the option of using a Tao brush for an endometrial biopsy in postmenopausal women.

Chapter 5

Discussion

Consideration of design, methods and recruitment

This study undertook a realistic evaluation of three contending outpatient methods for diagnostic evaluation of AUB, in terms of success of the technique itself and patient acceptability. At the same time it has been possible to conduct a rigorous evaluation of an alternative method of endometrial sampling (Tao brush).

The need for the present study was due to the lack of published evidence at the time of submission of the proposal as to the effectiveness of the various investigative modalities available for outpatient assessment when women present with AUB (including postmenopausal bleeding). The necessity for research that may determine the optimal investigation strategy in terms of effectiveness and cost-effectiveness has continued to be highlighted.[2] Hence, the current study was designed to enable a comparison of methods of outpatient endometrial evaluation taking into account the age variation in endometrial disease and the difference by menopausal status in the need for therapeutic management of the bleeding problem. The study has a broader remit than just efficiency in detection of abnormality and prevalence of adverse events. The study included assessment of the impact on management of care: patient acceptability, cost-effectiveness, and resource use and cost. The pragmatic study design ensured that the results should be generalisable to routine outpatient clinical practice. It is hoped that it will provide clinicians with background information to enable them to make an informed decision as to the most appropriate evaluation strategy for an individual woman. The accompanying cost-effectiveness data will be invaluable for those making decisions about gynaecology services to be provided. The study has been executed against a background of rapidly advancing technology over the time-frame of the investigation, especially in the context of hysteroscopic equipment, and the introductions of guidelines for the management of menstrual disorders in both primary and secondary care.[65,66] In addition, the Scottish Intercollegiate Guidelines Network (SIGN) guideline for management of postmenopausal bleeding has recently become available in Scotland.[24]

It is inevitable that evaluation of AUB will have a number of outcomes: detection of carcinoma, diagnosis of benign but clinically relevant uterine disease, treatment choice, clinician and/or patient reassurance, discharge and resource use. This study aimed to address diagnostic evaluation for the range of gynaecology clinic referrals for AUB, both premenopausal and postmenopausal. The study was not powered to replicate work on sensitivity and specificity for detection of endometrial cancer, which occurs predominantly in postmenopausal women. To do the latter would also require the gold-standard diagnosis of a pathological examination of the entire uterus, which would only be possible if all women proceeded directly to hysterectomy. This would mean an unacceptably narrow spectrum of women for study, so that the diagnostic performance observed would not necessarily be representative of the methods used in other clinic attendees. The design of the current study has placed a premium on the generalisability of the results, and hence all women attending the gynaecological service, with a complaint of AUB, were evaluated.

This study aimed to establish what investigation is most acceptable and most efficient in achieving diagnosis in the majority of referrals with an AUB complaint. Furthermore, it aimed to determine in premenopausal women which investigation strategy allows best management of their menstrual bleeding problem.

For benign but clinically relevant uterine disease, the estimation of the sensitivity and specificity of a diagnostic method is not possible because there is no gold-standard diagnosis that could be used for research. Nor is there complete consensus as to which abnormalities are clinically relevant. It is well known that not all fibroids cause heavy bleeding, and the clinical significance of polyps is not yet clear. In both conditions some occurrences may be associated with symptoms, but not necessarily all. Therefore, an investigation method that detected (and led to treatment of) structural variants of the uterus that were not of clinical relevance would waste resources, and sometimes subject women to needless risks. It is for this reason that resource use was reviewed for up to 2 years, to be able to calculate cost-effectiveness to

the end of investigation and to be able to compare accumulated costs between the different randomised options. The case-note review at 10 months and 24 months allowed an examination of not only the immediate costs and performance of evaluation method, but also accumulated healthcare costs, which may become inflated by an inconclusive or inadequately reassuring procedure.

In addition, the women were assessed by follow-up after 10 months and 2 years, so that it would be possible to compare the methods and their costs in terms of patient reassurance, satisfaction with the investigation undergone and impact on symptoms. Further, a unique aspect of this study has been the inclusion of an assessment of personality and psychological health at recruitment, important factors in complaining, and of the potential importance for explaining individual differences in anxiety, satisfaction about management or the extent to which reassurance is possible. The assessment of clinician's views has provided an insight into the extent to which involvement in a study such as this may be associated with a change in attitude to outpatient methods of endometrial evaluation.

The randomisation was undertaken separately within groups characterised by risk, to ensure ethical and clinical acceptability of randomisation options for the different risk groups. It is hoped that this approach has maximised statistical efficiency within the constraints of current best practice for clinical management. The nature of the study meant that blinding was impossible. The target for recruitment to the high-risk group was met. However, there was underrecruitment in the moderate- and low-risk groups (81.5% and 52.3% of the respective targets). This was partly due to a lower participation rate in the lower risk women, often because there was reluctance on the part of either or both of the low-risk woman and the referring clinician to allow randomisation to investigative option. For the low-risk group of women, under 40 years, the shortfall in recruitment occurred mainly because during the time-frame of the study there has been a change in gynaecologist strategies for the investigation of AUB, with many believing that no investigation is required. The lower participation rate could have been compensated for by approaching more women to participate, but a further change has been that over time fewer women eligible for the study were being referred to outpatients. This is likely to be because guidelines for the management of menorrhagia were made available for both primary and secondary care.[65,66] The consequences of this are that in the low-risk group the numbers studied are very small, and on the whole the differences in findings between subgroups cannot be interpreted with any confidence. Although this is disappointing, it is clear that women in this age group who are referred must be atypical of the usual run of women in that age group managed by their GPs. This is perhaps supported by their rating of their health as poorer than that of other women the same age, and by their acknowledgement of the extent to which they worry about their health.

Performance of investigative modality

The reasons why investigations cannot be undertaken fall into three general categories: medical, technical and patient related. There may be a very valid medical reason why a specific outpatient investigation cannot be undertaken, for example, the woman is too frail. Alternatively, there may be a technical reason for not undertaking the investigation, for example, the hysteroscope has not been sterilised and therefore cannot be used. Other reasons are more patient driven, for example, the woman withdraws from the study or fails to attend her appointment. Such failures may be influenced by current life circumstances, personality or previous exposure to investigations, and given randomisation, such factors should be evenly distributed across investigations. However, some reasons may also reflect the woman's attitude to a specific investigation, for example, as unpleasant or not crucial to diagnosis.

Ultrasound evaluation was completed in all cases where the woman attended, but there were cases where hysteroscopy was not possible owing to a failure to introduce the telescope or because this outpatient procedure was not deemed suitable for the woman. On an ITT basis adequate visualisation was achieved in 88% of high- and moderate-risk women having ultrasound and 77% undergoing hysteroscopy ($p = 0.002$).

An important aspect of the study was the ability of either hysteroscopy or ultrasound to detect an abnormality. Overall, in postmenopausal women relative to premenopausal women, fibroids were less often detected and polyps were more often detected. Across all risk groups (i.e. both postmenopausal and premenopausal women), ultrasound was demonstrated to be the best investigation for the detection of fibroids (32%

versus 13%, $p < 0.001$), while hysteroscopy performed better for the detection of endometrial polyps (12% versus 4%, $p < 0.001$). The significance of intrauterine polyps is not known and the majority are associated with benign pathology. A recent prospective survey of some 1500 UK consultant gynaecologists reported the variation in opinion concerning the role for both inpatient and outpatient treatment of polyps.[15] The majority of gynaecologists favoured inpatient removal, with half performing inpatient hysteroscopy to localise the polyp before blind removal. If there is considerable benefit then it might have been expected that postmenopausal women having hysteroscopy (and likely detection of polyps if present) would have had better outcome than those having ultrasound, and vice versa for the moderate-risk women in respect of detection of fibroids. At the 10-month follow-up postmenopausal women had similar rates of symptoms much improved, regardless of visualisation, and very slightly fewer women who had had hysteroscopy reported persisting symptoms (15% versus 17%) and failure to cure the problem (10% versus 14%). While persisting symptoms may suggest some undetected polyps in women having ultrasound, the effect had disappeared by the 24-month follow-up. For moderate-risk women, at 10 months the women who had had ultrasound were, if anything, very slightly worse off in terms of symptoms and cure, so the advantage for ultrasound in terms of detection of fibroids does not appear to convert into cure of the presenting problem or symptoms within 10 months. However, by 2 years' follow-up a slightly higher proportion of the ultrasound group was reporting symptoms much improved (61% versus 53%).

In postmenopausal women measurement by ultrasound of endometrial thickness greater than 4 mm may be an indication of endometrial cancer. Indeed, new guidelines have recently been introduced in Scotland that include commentary and guidance for clinical management of postmenopausal bleeding.[24] These recommend first line testing by TVUS and that, in postmenopausal women with prior probability of endometrial cancer less than 10%, a negative result using a threshold of 3 mm could, if patient and clinician were sufficiently reassured, avoid the need for biopsy. In the present study all postmenopausal women were randomised to biopsy in addition to visualisation. This allowed calculation of the sensitivity of ultrasound, which was 67% using the local threshold of 4 mm, but very imprecisely estimated, owing to the small number of cases of cancer among postmenopausal women randomised to TVUS ($n = 3$). If the 3 mm cut-off had been used the sensitivity would have been unchanged. Hysteroscopy can allow the observation of macroscopic changes to the endometrium suggestive of endometrial cancer. In this study the sensitivity for this in all women evaluated in this way was 20%, again with very wide confidence intervals.

Endometrial biopsy techniques are only of value when an adequate sample is obtained.[2] The first requirement for obtaining a sample is to be able to insert the sampling device. This study demonstrated that the Pipelle and Tao showed almost equal rates of failure of insertion (both 13%). This reflects the rates reported in other clinical trials.[18,67] The present study also considered individual characteristics that may influence the performance of the two endometrial sampling methods. Nulliparity was strongly associated with failure of insertion, whereas for parous women, the vast majority, the failure rate was only 8%. In this study the use of HRT, menopausal status and age had no significant influence on the success of insertion of either sampling method.

This study found that in high-risk women the Tao brush significantly outperformed the Pipelle sampler in terms of success in obtaining samples acceptable for pathological assessment. Considering all high-risk women randomised (ITT), the difference in percentage of adequate samples was 29 percentage points in favour of Tao brush (95% CI 21 to 37 percentage points). For the subset for whom insertion was possible, the corresponding difference was 35 percentage points (95% CI 26 to 45 percentage points). Among premenopausal women, deemed at moderate risk, the Pipelle was demonstrated to perform slightly better in terms of sample adequacy than the Tao brush (difference borderline significant, $p = 0.055$).

Among women for whom insertion was possible, the individual characteristics associated with obtaining an adequate sample differed for postmenopausal (high-risk) women and premenopausal (low- and moderate-risk) women. The percentage of adequate samples was generally high for premenopausal women for whom insertion was possible (86% Tao, 90% Pipelle), but in postmenopausal women differed by device (for Tao 88%, for Pipelle 52%). Therefore, menopausal status was important only for adequacy of samples using the Pipelle sampler. Within the postmenopausal women for whom insertion was possible, inadequate Pipelle samples were most common in those over 54 years of age (58% versus

33% and 36% for the younger two age groups, $p = 0.01$). For the Tao brush inadequate samples were also slightly more common in this oldest age group (18% versus 6% and 5%). These findings replicate other published studies which have demonstrated that age and menopausal status may be factors of importance for sample adequacy.[67,68] In the moderate-risk (premenopausal) women for whom insertion was possible, nulliparity was for both Tao brush and Pipelle samplers an important factor in the outcome 'inadequate sample' ($p < 0.001$ for both), in addition to its effect on insertion. These findings suggest that the use of Tao brush may be advantageous in the investigation of postmenopausal women, but not for moderate-risk women.

In both older premenopausal and, particularly, postmenopausal women, a difference was observed in success of obtaining a sample depending on whether the biopsy was undertaken blind or at the same time as hysteroscopy, with hysteroscopy having the advantage for both the Pipelle sampler (overall 70% versus 60%) and Tao brush (overall 80% versus 67%). A potential confounder to be considered is the experience of the operator collecting the sample. Given that this was a pragmatic trial there was no attempt to randomise to operator, nor would this have been feasible. It is possible that the clinicians undertaking hysteroscopy were more experienced at sampling, particularly with the newer Tao brush method. However, other technical factors should be considered. The potential influence of the introduction of gas or fluid (to facilitate hysteroscopy visualisation) on the success of the subsequent biopsy procedure is unknown. (In this study gas was used for insufflation.)

Patient acceptability

The large majority of women attending the clinic were reassured (87%) and glad that they had undergone their investigation (94%). Over three-quarters of women considered that their clinic visit was 'very' or 'extremely' worthwhile. No woman undergoing an ultrasound investigation had an adverse event and all adverse events encountered were minor. There was no significant difference in adverse events between hysteroscopy (12% of women) and biopsy (9% of women). Very few women (1%) described their experience of an ultrasound investigation as 'markedly unpleasant'. This was the case, however, for 18% women undergoing an endometrial biopsy and 16% of women having a hysteroscopy. The most negative reports for the investigation experience were from younger, low-risk women. Overall, 53% and 62% of women, respectively, reported some abdominal discomfort or vaginal bleeding following the investigative procedure. The women most likely to report such problems were those randomised to a hysteroscopy with biopsy or a biopsy alone, younger women and those with poorer self-reported general health.

Among young, low-risk women the report of the experience of collection of an endometrial sample with either a Pipelle sampler or a Tao brush was similar. Up to one-quarter of the small number of women in the group considered a biopsy to be markedly unpleasant and almost half described abdominal discomfort with the procedure. Among postmenopausal women (high risk) and premenopausal women (moderate risk) the study design involved biopsy by both methods, so a within-patient comparison of the experience with each of the biopsy techniques was possible (based on responses from only 63% of the women). There was a marginal but statistically significant preference for the Tao brush, with 33% stating that they preferred the Tao brush and 18% the Pipelle sampler.

At the time of investigation ultrasound was much more acceptable to women than hysteroscopy and biopsy, but hysteroscopy was not more unpleasant to women than biopsy. Women who had hysteroscopy were pleased to have had the investigation and women having this randomisation option were least likely to have wanted more investigation, whereas those having 'biopsy only' wished that they had had more and seemed unreassured. With time, hysteroscopy was more favourably viewed than ultrasound (10 months) or equivalent (24 months). There may be scope to reduce the unpleasantness of investigations by giving women more information about what to expect. It is noteworthy that on the whole women found biopsy alone more unpleasant than hysteroscopy. This may be due to the way in which they are prepared for hysteroscopy, compared with the more routine approach to taking a biopsy during the clinic appointment.

Economic evaluation

The definition of a full economic evaluation is the comparative analysis of alternative courses of action in terms of both their costs and their consequences. This trial lent itself well to economic evaluation as it was designed to reveal the relative

consequences of a range of comparators for the evaluation of women with AUB. The costing process for each alternative was patient based, with as much detailed information on direct NHS costs as possible recorded in the patient notes and coded using coding frames developed for the study. Where unit costs were not readily available they were calculated specifically for this study. The point at which an outcome was achieved was debated, as there was a number of complicating factors including the fact that treatments were often started before a satisfactory diagnosis was obtained. A decision was made to define the outcome end-point for the study participants principally as the point at which no further investigations took place. The premise was that a satisfactory diagnosis must have been reached if no further investigations were carried out.

The analysis was conducted separately for the women stratified into the high-risk, moderate-risk and low-risk groups, and whereas in the two lower risk groups the least cost-effective form of evaluation was found to be that which involved hysteroscopy, in the high-risk group it was marginally the more cost-effective method. In the moderate-risk group the most cost-effective method was blind biopsy, which is slightly surprising as biopsy is solely for exclusion of cancer, and seldom informs the management of abnormal menstrual bleeding. In the low-risk group the most cost-effective methods were blind biopsy or ultrasound. The incremental cost of not using a hysteroscopy combination method of evaluation in the high-risk group was calculated as £88 per patient, and of using more than blind biopsy for investigation of moderate-risk women as £212 per patient. However, the incremental cost of using the investigations other than ultrasound for the low-risk women was £452 per patient.

An estimation of the potential economic impact of selecting the most cost-effective means of evaluation for these women (rather than the least cost-effective alternative examined) shows that, where conservative figures from one hospital in Lothian are extrapolated to Scotland, a potential annual saving of less than £740,000 is possible for NHS Scotland. This is a relatively modest potential saving, representing 0.02% of the annual budget of £3377 million in 1999/2000 for Health Boards in Scotland to provide Hospital and Community Health Care. Therefore, it could be argued that other issues, such as clinician and women's preferences, should be weighted more heavily in helping to decide the evaluation method of choice.

At the time the study was designed the Tao brush was a new approach to endometrial evaluation, and it was not felt to be ethical to randomise high-risk women to biopsy by this device alone. Therefore, in the study design high- and moderate-risk women were assigned biopsy by both Pipelle sampler and Tao brush. This had benefits for increasing the power in terms of (paired) comparisons of performance and acceptability of these two samplers, but it did mean that they could not be compared in terms of cost-effectiveness, by the methodology outlined above. However, it was possible to estimate the additional costs accruing when a successful biopsy is not achieved, as might be the case for example with a Pipelle sampler, which has a high rate of inadequate samples in postmenopausal women. These estimated further costs were considerably in excess of the differential in unit costs of substituting a Tao brush biopsy for Pipelle, or the increase in investigative costs of adding a Tao brush biopsy to the investigations undertaken at the first appointment, if it can be judged by visual inspection that the Pipelle sampler has failed to provide an adequate sample.

Clinician attitudes

The assessment of local clinicians' views over the time-frame of the study provided an insight into clinicians' views as to the performance of the main outpatient methods of evaluation of AUB and of factors that may contribute to a change in attitude to these. A questionnaire was completed by the local clinicians (consultant gynaecologists and trainees) involved regularly in the care of women presenting to the clinic with AUB. The questionnaire addressed strategy preference and attitudes to the various outpatient methods of endometrial evaluation. Most clinicians expressed a lack of experience in the use of outpatient hysteroscopy or the Tao brush, but were comfortable with the use of a Pipelle sampler and the interpretation of reports on ultrasound investigations requested. These are aspects for consideration in the light of the findings of this study in the context of the potential for increased use of the Tao brush for endometrial sampling in postmenopausal women.

Although clinicians' judgements of general completion and success rates of the investigation options showed a wide range of performances, it should be remembered that the survey comprised a very small number of respondents, and some of the respondents may have been relatively

inexperienced in this area. The findings of the study were made available to all clinicians and three-quarters of clinicians indicated an interest in the option of using a Tao brush for an endometrial biopsy in postmenopausal women.

Integrated overview of decision as to investigation modality

The evaluation of diagnostic procedures necessarily involves a determination of the cost-effectiveness of the methods used. Although the randomisation options differed in terms of cost-effectiveness, this was not by such a margin that a benefit would be seen for the NHS budget in Scotland, if there were a universal shift to use of the most cost-effective evaluative option in each risk band. However, the study design, with pairing of Pipelle with Tao biopsies, was unable to compare empirically the cost-effectiveness of the Tao brush and Pipelle sampler. The additional costs that accrued for those where no study biopsy was adequate were calculated, and extrapolated to a hypothetical scenario where the only biopsy planned had been Pipelle but this had failed, and this fact could be seen by inspection. In such cases there would be a cost saving if a Tao brush was available to attempt a further biopsy by that method. This is likely to be particularly advantageous in postmenopausal women, both in terms of cost and in expediting diagnosis and management, and reducing worry and inconvenience to the woman.

In the present study Pipelle sampling was chosen because it is the most widely used method for endometrial biopsy. A recent systematic review demonstrated that the Pipelle sampler, compared with traditional D&C or hysterectomy, was an effective and accurate method for endometrial evaluation.[15] It is recognised that the Tao brush allows for both endometrial cytology and histology and hence may identify positive endometrial histology, potentially undetected with a Pipelle biopsy.[19] This is of particular relevance in the high-risk, postmenopausal, woman with an atrophic endometrium. The assessment of endometrial cytology, however, is recognised as a novel technique and cannot be conducted without specific training in the cytological assessment of both normal and abnormal endometrium. This aspect will inevitably incur additional costs in terms of the practical training of cytopathologists concerning the interpretation of endometrial tissue samples.

Diagnostic accuracy alone cannot be used to determine the value of a diagnostic test in a specific individual, since the relative importance of detection of polyps and fibroids is unknown, and the comparative sensitivities of the Tao brush and Pipelle sampler are imprecisely estimated in this study. However, it is noteworthy that the cancer cases missed by Pipelle were due to inadequate samples, and there is evidence regarding the adequacy of samples for a substantial number of women evaluated by both endometrial sampling methods. The superior performance by Tao found in this study, in terms of obtaining adequate samples, would mean, if extrapolated to a larger study evaluating, say, 2000 postmenopausal women and detecting 100 endometrial cancers (assuming an incidence of 5%), that over 20 cancers would be missed by Pipelle but detected by Tao.

The clinical decision regarding the most appropriate investigation will include issues of: index of suspicion of disease, the accuracy of the chosen biopsy method for detection of malignancy, the accuracy of either ultrasound or hysteroscopy for the detection of polyps or fibroids (sensitivity), and the frequency of a negative biopsy, ultrasound scan or hysteroscopy in a woman who has neither benign nor malignant uterine pathology (specificity). Unfortunately, the small clinician survey suggests that many clinicians do not know the performance criteria of the evaluation methods available to them. This is not surprising given the complexity of the data. It may be helpful to produce a pocket-sized card that summarises the essential features, and that can be readily updated as new data emerges. Almost all clinicians surveyed agreed that a clinical pathway would be welcome.

Since the ideal investigation or indeed combination of diagnostic evaluations is yet to be determined, there is a need to take into account the woman's preference, acceptability and cost-effectiveness.

Chapter 6

Recommendations for future research

For postmenopausal women exclusion of cancer is a main objective, so once the investigation has been completed discharge follows, but in the woman with abnormal menstrual bleeding, even if serious pathology is excluded, the original presenting symptoms require management. Therefore, in future research into the evaluation and management of AUB, postmenopausal women should be studied separately from premenopausal women with menstrual bleeding problems.

In premenopausal women attendance at an outpatient clinic did not help the majority of women with AUB, as about 60% reported that their symptoms were not much improved at 10 months. Patient perspectives are important, so research is needed to understand the relatively poor outcome for these patients, and to explore ways to integrate patient factors into optimising evaluation and treatment in these cases. In addition, the present study highlighted the need for research seeking to improve the dissemination of explanation and information to women being investigated for AUB.

The significance of benign pathologies in the premenopausal groups also requires clarification. Hysteroscopy and ultrasound visualisations have markedly disparate performance with respect to the detection of polyps and fibroids, with ultrasound detecting two to three times as many fibroids and hysteroscopy three to four times as many polyps. However, despite these marked differentials in diagnosis rates, there were minimal differences in judgements of 'symptoms improved'.

In postmenopausal women, the Tao brush biopsy is more effective in obtaining a sample adequate for pathological analysis compared with the Pipelle sampler biopsy. Given the performance of the Tao brush in this study there is now a need for a definitive study of the performance of Tao and Pipelle, taking into account inconclusive tests. In parallel, it would be essential to evaluate the implications for training of pathologists if there were to be increased use of the Tao brush biopsy technique in routine clinical practice.

In high- and moderate-risk women a higher proportion of adequate biopsy samples was obtained when biopsy was undertaken in conjunction with hysteroscopy. Research is required to ascertain whether this is due to the process (e.g. gas inflation of the uterus) or an operator effect, associated with the specific clinicians undertaking hysteroscopy.

There was relatively little difference in cost-effectiveness between different methods of evaluation of AUB. There is scope to make better use of patient factors to inform decisions as to the most efficient and acceptable method of investigation for an individual patient. Additional analyses, using data available as a result of this study, will contribute to this agenda. Future studies need to consider the impact of patient perspectives in detail.

Research is needed into useful and effective methods for the provision of succinct and simple information to clinicians regarding the performance of investigation methods *vis-à-vis* completion of tests, success in terms of visualisation or obtaining an endometrial sample, and detection of the various abnormalities. The possibility of developing a clinical pathway should also be considered.

Chapter 7
Conclusions

This study has highlighted the complexity of the investigation pathways travelled by women referred for AUB. Therefore, decision-making about investigation and understanding would be clarified if postmenopausal women were studied separately from premenopausal women.

The Tao brush was preferred by women and was superior in obtaining adequate samples, so it should be considered the method of choice for postmenopausal women, or at least be readily available as a back-up technique where Pipelle sampling has failed. The clinicians expressed interest in the Tao brush being made available for their use. Its introduction would have resource implications in terms of training of pathology staff.

Given the relatively small differences observed in cost-effectiveness between the methods of evaluation of AUB, there is justification for allowing other issues (such as clinician preferences and women's perspectives) to influence decisions as to the investigation method.

Acknowledgements

The other clinical study co-investigators, Dr Alistair Williams (Pathology, University of Edinburgh), Dr Sarah Chambers (Lothian University Hospitals Trust, Edinburgh) and Professor Andrew Calder (Obstetrics and Gynaecology, University of Edinburgh) were key to the design and successful conduct of the study, and we thank them for their very important contributions.

The authors and the above investigators thank the women who participated in the study, for sharing with us their feelings about their attendance at the outpatient gynaecology clinic, and for so diligently completing and returning the reports and questionnaires.

We also wish to express appreciation to Anne Douglas (who was involved in the early years of the study, including development of study materials, randomisation procedures and the initial setting up of data management systems) and Joan Kerr (who played a major part in recruitment, execution of the study and case-note review).

We gratefully acknowledge the contributions of our clinical colleagues, Jane Walker and Tom Fitzgerald, with ultrasound assessments of women in the study, and Sharon Cameron for her assistance with hysteroscopic evaluations during the latter part of this study. We thank Dr Joyce Willock for her advice on personality assessment and contributions to the development of questionnaires at the outset of the study, and for organising the initial focus groups. We also thank Kathy McGregor for advice at the start of the study on data collection for the economic evaluation.

We are grateful to Jane McCafferty for her enthusiasm and support of the study at the gynaecology outpatient clinic, and the many nursing and administration staff who helped in various ways, in particular to staff in Medical Records for facilitating the process of drawing notes for the case-note review. We would like to express appreciation of the help given by Audrey Burnside at a very difficult and busy time, by answering a stream of queries about suppliers and costs for the minutiae of clinic resourcing.

Finally, we would like to thank Dorothy Thom for her skilled data entry, not just for the many questionnaires but especially for the hugely complex case-note review data. Her insight and care contributed to the quality of the data available for analysis.

Contribution of the authors

HOD Critchley, P Warner, A Williams, S Chambers, A Calder were investigators and S Brechin, A Lee, J Guise, J Kerr, B Graham, A Douglas, J Walke, S Cameron, T Forster, J McCafferty, J Willock were members of the research team.

References

1. Gimpelson RJ, Rappold HO. A comparative study between panoramic hysteroscopy with directed biopsies and dilatation and curettage. *Am J Obstet Gynecol* 1988;**158**:489–92.
2. Clark TJ, Khan KS, Gupta JK. The diagnosis of intrauterine pathology in postmenopausal women: an evidence-based approach. *Rev Gynaecol Pract* 2002;**2**:109–16.
3. Creasman WT. Endometrial cancer: incidence, prognostic factors, diagnosis, and treatment. *Semin Oncol* 1997;**24**:S1–140.
4. Smith-Bindman R, Kerlikowske K, Feldstein VA, Subak L, Scheidler J, Segal M, *et al*. Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. *JAMA* 1998; **280**:1510–17.
5. Dijkhuizen FP, Brolmann HAM, Potters AE, Bongers MY, Heintz APM. The accuracy of transvaginal ultrasonography in the diagnosis of endometrial abnormalities. *Obstet Gynecol* 1996;**87**:345–9.
6. Vessey M, Clarke J, McKenzie I. Dilatation and curettage in young women. *Health Bull* 1979; **39**:59–62.
7. Lewis BV. Diagnostic dilatation and curettage in young women. *BMJ* 1993;**306**:225–6.
8. Penney GC. Endometrial sampling and D&C. Feedback report. Aberdeen: Gynaecology Audit Project in Scotland II. (Copies available from SPCERH Office, Aberdeen Maternity Hospital.)
9. Penney, G. Gynaecology Audit Project in Scotland II: Endometrial sampling and D & C – The audit loop closed: results from a second round of case note review. Edinburgh: Clinical Resource and Audit Group of SOHHD; 1997. p. 1.
10. Penney G, Vale L, Souter V, Templeton A. Endometrial assessment procedures: an audit of current practice in Scotland. *Hum Reprod* 1997; **12**:2041–5.
11. Ferry J, Farnsworth A, Webster M, Wren B. The efficacy of the pipelle endometrial biopsy in detecting endometrial carcinoma. *Austr N Z J Obstet Gynaecol* 1993;**33**:76–8.
12. Symonds I. Ultrasound, hysteroscopy and endometrial biopsy in the investigation of endometrial cancer. *Best Pract Res Clin Obstet Gynaecol* 2001;**15**:381–91.
13. Leather AT, Savvas M, Studd JW. Endometrial histology and bleeding patterns after 8 years of continuous combined estrogen and progesterone therapy in postmenopausal women. *Obstet Gynecol* 1991;**78**:1008–10.
14. 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**:1–55.
15. Clark TJ, Mann CH, Shah N, Song F, Khan KS, Gupta JK. Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: A systematic qualitative review. *Br J Obstet Gynaecol* 2002;**109**:313–21.
16. Clark TJ, Mann CH, Shah N, Song F, Khan KS, Gupta JK. Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial hyperplasia: a systematic quantitative review. *Acta Obstet Gynecol Scand* 2001;**80**:784–93.
17. Clark TJ, Voit D, Khan K, Hyde C, Song F, Gupta JK. Accuracy of hysteroscopy in the diagnosis of endometrial cancer: a systematic quantitative review. *JAMA* 2002;**288**:1610–21.
18. Gordon SJ, Westgate J. The incidence of failed pipelle sampling in a general outpatient clinic. *Aust N Z J Obstet Gynaecol* 1999;**39**:115–18.
19. Maksem J, Knesel E. Liquid fixation of endometrial brush cytology ensures a well-preserved, representative cell sample with frequent tissue correlation. *Diagn Cytopathol* 1996;**14**:367–73.
20. Kremer C, Duffy S, Moroney M. Patient satisfaction with outpatient hysteroscopy versus day case hysteroscopy: randomised, controlled trial. *BMJ* 2000;**320**:279–82.
21. Weiser F, Tempfer C, Kurtz C, Nagele F. Hysteroscopy in 2001: a comprehensive review. *Acta Obstet Gynecol Scand* 2001;**80**:773–83.
22. 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.
23. Gupta JK, Chien PFW, 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**:799–816.

24. Scottish Intercollegiate Guidelines Network. Investigation of postmenopausal bleeding. Edinburgh: SIGN; 2002. p. 61.

25. Cameron ST, Walker JJ, Champeney R, Critchley HOD. Comparison of transvaginal ultrasound scan, saline infusion hysterography and hysteroscopy for investigating postmenopausal bleeding and unscheduled bleeding on hormone replacement therapy. *Int J Obstet Gynaecol* 2000; Abstract Series 13.

26. Bradley LD, Falcone T, Magen AB. Radiographic imaging techniques for the diagnosis of abnormal uterine bleeding. *Obstet Gynecol Clin North Am* 2000;**27**:245–6.

27. Bradley LD, Pasqualotto EB, Price LL, Margossian H. Hysteroscopic management of endometrial polyps. *Obstet Gynecol* 2000;**4**(1):S23.

28. Emanual M, Verdel MJ, Wamsteker K, Lammes F. A prospective comparison of transvaginal ultrasonography and diagnostic hysteroscopy in the evaluation of patients with abnormal uterine bleeding: clinical implications. *Am J Obstet Gynecol* 1995;**172**:547–52.

29. Spencer CP, Cooper AJ, Whitehead MI. Management of abnormal bleeding in women receiving hormone replacement therapy. *BMJ* 1997;**315**:37–42.

30. Zethraeus N, Johannesson M, Henriksson P, Strand RT. The impact of hormone replacement therapy on quality of life and willingness to pay. *Br J Obstet Gynaecol* 1997;**104**:1191–5.

31. Nagele F, O'Connor H, Davies A, Badawy A, Mohammed H, Magos AL. 2500 outpatient diagnostic hysteroscopies. *Obstet Gynecol* 1996; **88**:87–92.

32. Van den Bosch T, Vandendael A, Van Schoubroeck D, Wranz PA, Lombard CJ. Combining vaginal ultrasonography and office endometrial sampling in the diagnosis of endometrial disease in postmenopausal women. *Obstet Gynecol* 1995; **85**:349–52.

33. Chambers JJ, Chambers SK. Endometrial sampling: When? Where? Why? With what? *Clin Obstet Gynecol* 1992;**35**:28–9.

34. Stovall TG, Ling FW, Morgan PL. A prospective randomised comparison of the pipelle endometrial sampling device with the Novak curette. *Am J Obstet Gynecol* 1991;**165**:1287–90.

35. Bistoletti P, Hjerpe A. Routine use of endometrial cytology in clinical practice. *Acta Cytol* 1993; **37**:867–70.

36. Tao LC. *Cytopathology of the endometrium: direct endometrial sampling*. Chicago: ASCP Press; 1993.

37. Karlsson B, Granberg S, Wikland M, Ylöstalo P, Torvid K, Marsal K, Valentin L. Transvaginal ultrasonography of the endometrium in women with postmenopausal bleeding – a Nordic multicenter study. *Am J Obstet Gynecol* 1995; **172**:1488–94.

38. Ferazzi E, Torri V, Trio D. Sonographic endometrial thickness: a useful test to predict atrophy in patients with postmenopausal bleeding. An Italian multicenter study. *Ultrasound Obstet Gynecol* 1996;**7**:315–21.

39. Huhta CA, Hall DA, McCarthy K. Endometrial polyps, hyperplasia and carcinoma in postmenopausal women: differentiation with endovaginal sonography. *Radiology* 1994; **191**:755–8.

40. Shushan A, Protopapas A, Hart R, Magos A. Diagnostic and therapeutic advantages of hysteroscopic surgery in management of intrauterine lesions in postmenopausal women. *J Am Assoc Gynecol Laparosc* 2001;**8**:87–91.

41. Revel A, Shushan A. Investigation of the infertile couple. Hysteroscopy with endometrial biopsy is the gold standard investigation for abnormal uterine bleeding. Debate. *Hum Reprod* 2002; **17**:1947–9.

42. Gebauer G, Hafner A, Siebzehnrubl E, Lang N. Role of hysteroscopy in detection and extraction of endometrial polyps: results of a prospective study. *Am J Obstet Gynecol* 2001;**184**:59–63.

43. Tahir MM, Bigrigg MA, Browning JJ, Brookes ST, Smith PA. A randomised controlled trial comparing transvaginal ultrasound, outpatient hysteroscopy and endometrial biopsy with inpatient hysteroscopy and curettage. *Br J Obstet Gynaecol* 1999;**106**:1259–64.

44. Garuti G, Sambruni I, Colonnelli M, Luerti M. Accuracy of hysteroscopy in predicting histopathology of endometrium in 1500 women. *J Am Assoc Gynecol Laparosc* 2001;**8**:207–13.

45. Lo KW, Yuen PM. The role of outpatient diagnostic hysteroscopy in identifying anatomic pathology and histopathology in the endometrial cavity. *J Am Assoc Gynecol Laparosc* 2000;**7**:381–5.

46. Indman PD. Abnormal uterine bleeding: accuracy of vaginal probe ultrasound in predicting abnormal hysteroscopic findings. *J Reprod Med* 1995;**40**:545–8.

47. Wood C, Hurley VA, Leoni M. The value of ultrasound in the management of menorrhagia. *Austr N Z J Obstet Gynaecol* 1993;**33**:198–200.

48. Nagele F, Badawy A, Magos AL. A comprehensive one-stop menstrual problem clinic for the diagnosis and management of abnormal uterine bleeding. *Br J Obstet Gynaecol* 1996;**103**:851.

49. Nagele F, Davies A, O'Connor H, Mohammed H, Baskett TF, Magos AL. Hysteroscopy in women

with abnormal uterine bleeding on hormone replacement therapy: a comparison with postmenopausal bleeding. *Fertil Steril* 1996; **65**:1145–50.

50. Pal L, Lapensee L, Toth TL, Isaacson KB. Comparison of office hysteroscopy, transvaginal ultrasonography and endometrial biopsy in evaluation of abnormal uterine bleeding. *Journal for the Society of Laparoendoscopic Surgery* 1997; **1**:125–30.

51. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. 2nd ed. Oxford: Oxford University Press; 1997.

52. HMSO. *'Designed to care': renewing the National Health Service in Scotland*. CM 3811. Edinburgh: The Stationery Office; 1997.

53. Costa P, McCrae R. *The NEO personality inventory manual*. Odessa, FL: Psychological Assessment Resources; 1985.

54. Goldberg DP, Hillier VF. A scaled version of the general health questionnaire. *Psychosom Med* 1979;**9**:139–45.

55. Mayou R, Bass C, Sharpe M. Overview of epidemiology, classification, and aetiology. In Mayou R, Bass C, Sharpe M, editors. *Treatment of functional somatic symptoms*. Oxford: Oxford University Press; 1995. pp. 42–65.

56. Aleong J, Bartlett DE. Improved groups for calculating sample sizes when comparing two independent binominal distributions. *Biometrics* 1979;**35**:875–81.

57. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to BMJ. *BMJ* 1996;**313**:275–83.

58. Graves N, Walker D, Raine R, Hutchings A, Roberts J. Cost data for individual patients included in clinical studies: no amount of statistical analysis can compensate for inadequate costing methods. *Health Econ* 2002;**11**:735–9.

59. Graham B. British Medical Association Health Policy and Economic Research Unit. 2000.

60. Personal Social Services Research Unit. *Unit costs of health and social care*. Canterbury: University of Kent; 2001.

61. Scottish Health Service Costs 2000–2001. Edinburgh: Information and Statistics Division; 2001.

62. STARD Group. Towards complete and accurate reporting of studies of studies on diagnostic accuracy: the STARD initiative. [Online] Test version November 2003. Amsterdam. Hans Reitsma. Available from: http://www.consort-statement.org/stardstatement.htm [accessed 18 Nov. 2003].

63. Altman DG, Schulz KF, Moher D, Davidoff F, Elbourne D. The revised CONSORT statement for reporting randomised trials: explanation and elaboration. *Ann Intern Med* 2001;**134**: 663–94.

64. Wilson EB. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 1927;**22**: 209–12.

65. Royal College of Obstetricians and Gynaecologists. *The initial management of menorrhagia*. London: RCOG Press; 1998.

66. Royal College of Obstetricians and Gynaecologists. *The management of menorrhagia in secondary care*. London: RCOG Press; 1999.

67. Batool T, Reginald PW, Hughes JH. Outpatient Pipelle endometrial biopsy in the investigation of postmenopausal bleeding. *Br J Obstet Gynaecol* 1994;**101**:545–6.

68. Bakour SH, Khan KS, Gupta JK. Controlled analysis of factors associated with insufficient sample on outpatient endometrial biopsy. *Br J Obstet Gynaecol* 2000;**107**:1312–14.

69. Department of Health. *National schedule of reference costs for England and Wales*. London: Department of Health; 2001.

70. Royal Pharmaceutical Society of Great Britain. *British National Formulary*. 40th ed. London: BMJ Books; 2000.

Appendix 1

Clinician-completed forms

ELIGIBILITY/RECRUITMENT SLIP

Date : |_|_|/|_|_|/|_|_|_|_|
dd mm yyyy

Patient name & address label

Patient eligible: ☐ confirmed,
- Referral complaint "abnormal bleeding" confirmed
- Patient NOT pregnant
- Patient has NO problems reading/writing English

Confirmation of group

Reproductive/Risk status

A ☐ Postmenopausal
B ☐ Premenopausal, age >=40; or age <40 + other factors
C ☐ Premenopausal, age <40 and no other factors

NB: Patients in group C ('low risk') have a 40% chance of being randomised to the "No evaluation" option, which is considered appropriate management for this group of women.

Other factors (age < 40)

☐ Polycystic ovarian syndrome
☐ Prior use of unopposed oestrogens/tamoxifen
☐ Obesity
☐ Diabetes
☐ Family history of endometrial ca

Clinician agreement to patient's participation in study:

☐ Yes
☐ No, (specify reason): _____

Clinician's evaluation *preference(s)*: (tick one or more boxes)

☐ No evaluation
☐ Biopsy
☐ Hysteroscopy
☐ Ultrasound
☐ Other: _____

Initials: _____

PATIENT

Date : ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

Name: _____

Address: _____

_____

_____

Postcode: _____ _____

Trial No : ☐☐☐☐

Date of birth : ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

Hosp No : ☐☐☐☐☐☐–☐☐☐☐

Managing clinician : _____

Informed consent : ☐ obtained, form signed

Presenting complaint(s):

- ☐ PMB
- ☐ Heavy periods
- ☐ PCB
- ☐ IMB
- ☐ Irreg. bleeding
- ☐ Other: _____

Previous menstrual complaint(s):

- ☐ PMB
- ☐ Heavy periods
- ☐ PCB
- ☐ IMB
- ☐ Irreg. bleeding
- ☐ Other: _____

	No	Yes
On oral contraception:	☐	☐
On HRT:	☐	☐
Sterilised:	☐	☐

LMP : ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

Days of loss (average): ☐☐

Cycle length (average): ☐☐

Reproductive/Risk status

- A ☐ Postmenopausal
- B ☐ Premenopausal, age >=40; or age <40 + other factors
- C ☐ Premenopausal, age <40 and no other factors

Other factors (age < 40)

- ☐ Polycystic ovarian syndrome
- ☐ Prior use of unopposed oestrogens/tamoxifen
- ☐ Obesity
- ☐ Diabetes
- ☐ Family history of endometrial ca

Clinicians evaluation preference(s):
(tick one or more boxes)

- ☐ No evaluation
- ☐ Biopsy
- ☐ Hysteroscopy
- ☐ Ultrasound
- ☐ Other: _____

Assigned **evaluation method:**

randomisation option label

Questionnaire completed (attached)

Well-being questionnaire (GHQ)	☐	
Personality factors questionnaire (NEO)	☐	
Patient's attitude to health	☐	
Patient's satisfaction questionnaire	☐	
Patient's experience of investigation(*s*)	☐	No eval. ☐ US ☐ H+B ☐ Biopsy
Patient reassurance questionnaire	☐	

Next outpatient appointment (if): ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

Patient hospitalised: ☐ No ☐ Yes

Comments: _____

_____

Initials: _____

Study: OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

ULTRASOUND

Date : ☐☐/☐☐/☐☐☐☐ (dd / mm / yyyy)

Trial number: ☐☐☐☐

Name: _____

Hosp No: ☐☐☐☐☐☐–☐☐☐☐

Clinic/Site: ☐ RIE ☐ GOPD ☐ OUTREACH

Sonographer: ☐ JW ☐ SC ☐ Other: _____

Machine #: _____

Ultrasound number: ☐☐☐☐/☐☐☐☐☐

Sonographer Technique: ☐ transvaginal ☐ transabdominal

Uterine size: ☐ normal
☐ increased → ____.___ cm x ____.___ cm x ____.___ cm

View of endometrium obtained: ☐ good ☐ difficult ☐ not seen

Endometrial thickness: ____. ___ mm

Endometrial morphology: ☐ normal
- ☐ early follicular
- ☐ late follicular
- ☐ secretory
- ☐ atrophic

☐ abnormal
- ☐ homogeneous
- ☐ containing cystic spaces
- ☐ heterogeneous
- ☐ polyp
- ☐ filling defect

Myometrial invasion:
- ☐ No
- ☐ Yes
- ☐ not relevant

Fluid in uterine cavity:
- ☐ No
- ☐ Yes

Uterine fibroids:
- ☐ No
- ☐ Yes →

Number	Site	Size
☐ 1 ☐ 2 ☐ multiple	☐ involving cavity ☐ not involving cavity	____cm

Lt Ovary
- ☐ not seen → Adnexal mass: ☐ No ☐ Yes → **Description**: ☐ simple cyst ☐ loculated cyst ☐ complex ☐ solid
- ☐ seen → Size: _____ ml → Morphology: ☐ normal ☐ abnormal → **Description**: ☐ simple cyst ☐ loculated cyst ☐ complex ☐ solid

Rt Ovary
- ☐ not seen → Adnexal mass: ☐ No ☐ Yes → **Description**: ☐ simple cyst ☐ loculated cyst ☐ complex ☐ solid
- ☐ seen → Size: _____ ml → Morphology: ☐ normal ☐ abnormal → **Description**: ☐ simple cyst ☐ loculated cyst ☐ complex ☐ solid

Diagnosis: ☐ normal
☐ abnormal (specify below)
- ☐ Premenopausal, structural abnormality
- ☐ Premenopausal, other: _____
- ☐ Postmenopausal, thickened
- ☐ Postmenopausal, other: _____

Other findings/comments: _____

_____

Initials: _____

Time started: ☐☐:☐☐ (24-hour clock)
Time finished: ☐☐:☐☐ (24-hour clock)
(do not include time for filling in study form(s))

Additional ultrasound appointment required: ☐ No ☐ Yes

Any adverse events?:
☐ No ☐ Yes (specify):

ENDOMETRIAL BIOPSY

Date : ☐☐/☐☐/☐☐☐☐ (dd / mm / yyyy)

Study: OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

Trial No: ☐☐☐☐

Name: _____

Hosp No: ☐☐☐☐☐☐-☐☐☐☐

Clinic/Site: ☐ RIE ☐ GOPD ☐ OUTREACH ☐ 'ONESTOP'

Clinician's initials: _____

Unit No.: _____

	Yes	No
Brush biopsy	☐	☐
Pipelle biopsy	☐	☐

If both, which first: ☐ brush ☐ pipelle

Sample quality:
- ☐ Good sample
- ☐ Scanty curettings
- ☐ No sample
- ☐ Failed insertion

Number of attempts: ☐☐

Findings:

Cervix: ☐ Nulliparous ☐ Parous

Cervix: ☐ Normal ☐ Abnormal →
- ☐ ectopy
- ☐ Cervicitis
- ☐ Nabothian follicles
- ☐ suspicious
- ☐ Polyp
- ☐ Other (specify): _____

Time started: ☐☐:☐☐ (24-hour clock)

Time finished: ☐☐:☐☐ (24-hour clock)

(do not include time for filling in study form(s))

Additional investigation required: ☐ No ☐ Yes

Adverse events:
- ☐ No
- ☐ Cervical shock
- ☐ Perforation
- ☐ Patient distressed
- ☐ Infection
- ☐ Other (specify):

Other findings/comments: _____

_____

Initials: _____

HYSTEROSCOPY + BIOPSY

Date : ☐☐/☐☐/☐☐☐☐ (dd / mm / yyyy)

Trial number: ☐☐☐☐

Name: _____

Hosp No: ☐☐☐☐☐☐–☐☐☐☐

Clinic/Site: ☐ RIE ☐ GOPD ☐ OUTREACH ☐ 'ONESTOP'

Clinician's initials: _____

Anaesthesia/Medication used: ☐ No ☐ Yes

if yes, type + dose: _____

_____

Patient watched video: ☐ No ☐ Yes

Cervical dilatation: ☐ No ☐ Yes

	No	Yes
IUCD in situ:	☐	☐
IUS in situ:	☐	☐

Endocervical canal: ☐ normal ☐ abnormal

Uterine body: ☐ normal ☐ abnormal → ☐ Polyp ☐ Fibroid ☐ Adhesion ☐ Atrophic ☐ Possible ca ☐ Possible hyperplasia

Ostium:	normal	abnormal
Left	☐	☐
Right	☐	☐

Lateral Wall:	normal	abnormal
Left	☐	☐
Right	☐	☐

Hysteroscopy completed: ☐ Yes ☐ No (specify reason):

Number of attempts (hysteroscopy): ☐☐

	Yes	No
Brush biopsy	☐	☐
Pipelle biopsy	☐	☐

If both, which first: ☐ brush ☐ pipelle

Sample quality: ☐ Good sample ☐ Scanty curettings ☐ No sample ☐ Failed insertion

Number of attempts (biopsy): ☐☐

Findings:

Cervix: ☐ Nulliparous ☐ Parous

Cervix: ☐ Normal ☐ Abnormal → ☐ Ectopy ☐ Cervicitis ☐ Nabothian follicles ☐ Suspicious ☐ Polyp ☐ Other (specify): _____

Diagnosis: ☐ Normal ☐ Abnormality → ☐ non-benign causing symptoms requiring treatment ☐ uncertain

Other findings/comments: _____

_____

_____

Time started : ☐☐:☐☐ (24-hour clock)

Time finished: ☐☐:☐☐ (24-hour clock)

(do not include time for filling in study form(s))

Adverse events:
- ☐ No
- ☐ Cervical shock
- ☐ Perforation
- ☐ Patient distressed
- ☐ Infection
- ☐ Other (specify):

Initials: _____

Appendix 2

Pathologist-completed forms

PATHOLOGY – BRUSH BIOPSY

Date : ☐☐/☐☐/☐☐☐☐ (dd / mm / yyyy)

Trial No: ☐☐☐☐

Name: _____

Hosp No: ☐☐☐☐☐☐–☐☐☐☐

Pathologist: ☐ AW ☐ Other: _____

UB No: ☐☐☐☐☐☐/☐☐

Source hospital/clinic: _____

Confirm use of Tao brush sampler: ☐

If both brush AND pipelle, specify order of analysis: ☐ brush first ☐ brush second

Adequacy of specimen:
- ☐ adequate
- ☐ barely adequate
- ☐ inadequate (insufficient)
- ☐ inadequate (technical artefact)
- ☐ inadequate (other) – specify: _____
- ☐ not applicable
- ☐ not received
- ☐ unknown

Diagnosis:
- ☐ inadequate
- ☐ atrophic
- ☐ inactive
- ☐ proliferative
- ☐ early secretory
- ☐ mid-secretory
- ☐ late secretory
- ☐ menstrual
- ☐ simple (cystic) hyperplasia
- ☐ complex hyperplasia, non-atypical
- ☐ complex hyperplasia, atypical
- ☐ adenocarcinoma, grade 1
- ☐ adenocarcinoma, grade 2
- ☐ adenocarcinoma, grade 3
- ☐ drug effect
- ☐ polyp
- ☐ not applicable
- ☐ not known
- ☐ other – specify: _____

Accompanying specimen: ☐ No ☐ Yes → **UB No:** ☐☐☐☐☐☐–☐☐

Diagnosis of accompanying specimen: _____

	Time started	Time finished
technical:	☐☐:☐☐ (24-hour clock)	☐☐:☐☐ (24-hour clock)
medical:	☐☐:☐☐ (24-hour clock)	☐☐:☐☐ (24-hour clock)

(do not include time for filling in study form(s))

Additional specimen required: ☐ No ☐ Yes

Other findings/comments: _____

_____

Initials: _____

PATHOLOGY – PIPELLE BIOPSY

Date : ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

Trial No: ☐☐☐☐

Name: _____

Hosp No: ☐☐☐☐☐☐–☐☐☐☐

Pathologist: ☐ AW ☐ Other: _____

UB No: ☐☐☐☐☐☐/☐☐

Source hospital/clinic: _____

Confirm use of pipelle sampler: ☐

If both brush AND pipelle, specify order of analysis: ☐ pipelle first ☐ pipelle second

Adequacy of specimen:
- ☐ adequate
- ☐ barely adequate
- ☐ inadequate (insufficient)
- ☐ inadequate (technical artefact)
- ☐ inadequate (other) – specify: _____
- ☐ not applicable
- ☐ not received
- ☐ unknown

Diagnosis:
- ☐ inadequate
- ☐ atrophic
- ☐ inactive
- ☐ proliferative
- ☐ early secretory
- ☐ mid-secretory
- ☐ late secretory
- ☐ menstrual
- ☐ simple (cystic) hyperplasia
- ☐ complex hyperplasia, non-atypical
- ☐ complex hyperplasia, atypical
- ☐ adenocarcinoma, grade 1
- ☐ adenocarcinoma, grade 2
- ☐ adenocarcinoma, grade 3
- ☐ drug effect
- ☐ polyp
- ☐ not applicable
- ☐ not known
- ☐ other – specify: _____

Accompanying specimen: ☐ No ☐ Yes → **UB No:** ☐☐☐☐☐☐–☐☐

Diagnosis of accompanying specimen: _____

	Time started	**Time finished**
technical:	☐☐:☐☐ (24-hour clock)	☐☐:☐☐ (24-hour clock)
medical:	☐☐:☐☐ (24-hour clock)	☐☐:☐☐ (24-hour clock)

(do not include time for filling in study form(s))

Additional specimen required: ☐ No ☐ Yes

Other findings/comments: _____

_____

Initials: _____

Study: OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

Appendix 3

Focus groups on consultation for bleeding problems

Aims

To elucidate:

- the issues for women consulting for abnormal bleeding
- their experiences of clinic organisation and the various procedures
- their perceptions regarding treatment and investigations.

Methodology

The consultant gynaecologist on the study asked a number of women attending the hospital gynaecological clinic for problem bleeding whether they would consider participating in a focus group. If they agreed to be contacted, their details were given to the researchers and they were then contacted by J Willock (JW) with an invitation to a specific group. Three hospital focus groups were arranged and led by JW, outside clinic hours, to discuss consultations for bleeding problems. Consent to participate was implicit in attendance.

The women who participated were predominantly in the middle to high socio-economic groupings, so the scope of the study was widened. Two further focus groups were organised in a community project group targeted at women in a socio-economically deprived area, and one in a Bengali health and welfare group set up for and run by south Asian women. In these groups, the women's experience of seeing a doctor for gynaecological problems was less recent, and the problems for which they consulted were more diverse. Approval was sought from the group facilitators, who explained the aims of the focus group and asked women whether they would like to participate at a particular session. Consent was implicit in attendance.

In total, eight focus groups of around 1 hour's duration were held, with a total of 24 women whose ages ranged from 30 to 50 years, 13 of whom were current hospital clinic patients.

The women in each group were initially asked to describe their experiences when attending a doctor for gynaecological or period problems. A researcher then used prompts to encourage them to relate their experiences of the problem, and their views on patient–doctor communication and on clinic organisation and procedures. The entire discussion was audiotaped and transcribed.

Analysis and main themes

A qualitative, content analysis of the transcriptions was carried out independently by two researchers. Three themes were examined:

- subjective report of the problem and GPs' reactions to this report
- the exchange of information in the consultation
- clinic organisation and procedures.

These were examined from two points of view: that of the women currently attending a hospital clinic, and that of the women in community groups. These analyses were then compared and discussed, and the original material was revisited, before the final analysis.

A number of issues emerged in the focus groups. With regard to perceptions of problem, when prompted about worry with respect to underlying causes of symptoms, this was initially dismissed, often quite categorically. Yet it was frequently the case that later in the discussion it would be stated that the outcome of the consultation was 'relief that nothing serious was wrong' or it was admitted that the whole process had been very worrying. It was also clear from discussion that there were substantial interindividual variations in time to consultation from onset of symptoms, and some suggestion that this may be related to socio-economic status. The nature of the disruption that symptoms caused to daily life also differed across the groups, with some highlighting disruption to work and social life, and others embarrassment and isolation.

The key issue that emerged from this study was the need for better communication, across a range of aspects: waiting times, gender of clinician, student presence, medical investigations and levels of reassurance. This problem was clearly acute for women for whom English was not a first or fluent language, often exacerbated by cultural factors that made discussion of gynaecological matters uncomfortable, but even those whose first language was English recounted problems with understanding technical language used by clinicians. This meant that explanations were often not understood, and that many women were inhibited from asking questions in case they seemed foolish.

Conclusions

Questionnaires to be developed for the study trial should address women's feelings about their symptoms, their anxiety about their symptoms and any investigations required, the prior information they had received about the clinic, and the extent to which they had received explanations from clinic staff about their symptoms or investigations.

Postscript

The questionnaires developed are presented in Appendices 4 and 5. The focus group analysis is being written up for publication as a research paper.

Appendix 4

Forms and questionnaires completed by the woman around the time of recruitment and investigation

• 22/2/99 **Date:** dd/mm/yyyy **Trial No:** ____

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

Date of birth: dd/mm/yyyy

HEALTH QUESTIONNAIRE

(1) Generally speaking, I would say my **current state of health** is:
- ☐ Very good
- ☐ Average
- ☐ Poor
- ☐ Very poor

(2) My general health **compared to other women** about my age is:
- ☐ Worse than most
- ☐ About the same
- ☐ Better than most

	Strongly disagree	Disagree	Agree	Strongly agree
(3) It is important to **check your body for signs** that something is wrong.	☐	☐	☐	☐
(4) Menstrual bleeding problems such as I have are to be **expected** in women of my age.	☐	☐	☐	☐
(5) I believe it should be possible for my problem to be **sorted out** quickly.	☐	☐	☐	☐
(6) Problems with bleeding are very **worrying**.	☐	☐	☐	☐
(7) If I have a medical problem I like the **doctor to decide** how to deal with it.	☐	☐	☐	☐
(8) I try in non-medical ways to **improve/maintain** my health (e.g. exercise, health foods).	☐	☐	☐	☐
(9) I like to **be told as much as possible** about my condition by my doctor.	☐	☐	☐	☐
(10) I try to **find out as much as possible** about my condition for **myself**, from magazines, radio, books, friends etc.	☐	☐	☐	☐
(11) I like a doctor to give me **choice** between the possible medical treatments or tests I may need.	☐	☐	☐	☐

(12) I feel my **symptoms** of bleeding may be due to:

*(Please tick one answer for **each** part)*	Not very likely	Slightly likely	Fairly likely	Very likely
ⓐ Fibroids	☐	☐	☐	☐
ⓑ Endometriosis	☐	☐	☐	☐
ⓒ Cancer	☐	☐	☐	☐
ⓓ Hormones	☐	☐	☐	☐
ⓔ Stress	☐	☐	☐	☐
ⓕ Family history	☐	☐	☐	☐

PTO

(13) If you were to need an investigation for your gynaecological problems, **how would you feel** …

*(Please tick one answer for **each** part)*

	No problem	Slightly unhappy	Unhappy	Very unhappy
(a) …if it required an investigation in **outpatients** of the inside of your womb?	☐	☐	☐	☐
(b) …if it required you to go to the hospital **operating theatre**?	☐	☐	☐	☐
(c) …if it required a **general anaesthetic**?	☐	☐	☐	☐

	No	Yes, a bit	Yes, very much
(14) Do you **worry** about your health?	☐	☐	☐
(15) Do you think there is something **seriously wrong** with your body?	☐	☐	☐
(16) Does your **family** have a **history of illness**?	☐	☐	☐
(17) Do you find that you are **aware** of various **things happening in your body**?	☐	☐	☐
(18) Are you **bothered by aches and pains**?	☐	☐	☐
(19) Are you **more sensitive to pain** than other people?	☐	☐	☐
(20) Do you find that you are **bothered by many different symptoms**?	☐	☐	☐

(21) Did you **fill in** this questionnaire at the clinic or at home? ☐ Clinic ☐ Home

Thank you very much for your help!

• 22/2/99 **Date:** |_|_|/|_|_|/|_|_|_|_| (dd mm yyyy) **Trial No:** |_|_|_|_|

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

Thank you, your help with this research is greatly appreciated. Remember, your answers will be kept quite confidential and will not be included in your medical notes. Some questions need you to write in an answer or a number, in most cases you will only need to tick a box. If questions have parts a,b,c…, please answer every part.

Date of birth: |_|_|/|_|_|/|_|_|_|_| (dd mm yyyy)

ABOUT THIS CLINIC VISIT

1 What are the **symptoms** that brought you to this clinic? _____
(Please write in your own words)
_____

2 There are different **reasons people go to a doctor with symptoms**. Please say how much the following reasons apply to your visit to the doctor. *(Please tick one answer for **each** part)*

I went to the doctor about my symptoms because...	Not at all	Slightly true	True	Very much so
ⓐ…*I have heard that is what women should do if they notice changes.*	☐	☐	☐	☐
ⓑ…*I was worried the symptoms were a sign of disease.*	☐	☐	☐	☐
ⓒ…*I want the symptoms to stop, whatever their cause.*	☐	☐	☐	☐

3 How do the symptoms that have brought you to this clinic make you **feel**?
*(Please tick a box in **each** of the first four rows)*

	Not at all	Not very	Quite	Very	Extremely
a *Anxious*	☐	☐	☐	☐	☐
b *Angry*	☐	☐	☐	☐	☐
c *Fed up*	☐	☐	☐	☐	☐
d *Embarrassed*	☐	☐	☐	☐	☐
e *Other* *optional*	☐	☐	☐	☐	☐

please specify: _____

4 **Before** coming here today, did you realise you were **likely to have to have an internal examination**? No ☐ Yes ☐

5 This is a teaching hospital, so often students will be present during consultations and investigations. How do you **feel about students being present** when you see the doctor?

Very unhappy	Unhappy	Would prefer not to have students	Do not mind	Quite happy
☐	☐	☐	☐	☐

6 Were there any **students present** at your visit today? No ☐ Yes ☐ → If yes, **how many**? |_|_|

7 Did your GP give you **information** about this clinic visit and **what to expect**?

☐ No information given
☐ Information given – but not enough
☐ Information given – sufficient

PTO

(8) Were you given any **written information before today** about what to expect of your clinic visit? No ☐ Yes ☐

If yes, was this: ☐ from your GP?
☐ sent with your appointment?
☐ both?

(9) Have you ever had any **investigations** for vaginal bleeding **before**? *(Please tick one answer for **each** part)*

	No	Yes	Don't know
(a) *Biopsy* (sample of tissue taken from inside your womb)	☐	☐	☐
(b) *Vaginal ultrasound* (ultrasound scan of your womb)	☐	☐	☐
(c) *Hysteroscopy* (viewing the inside of the womb through a lens on a monitor)	☐	☐	☐
(d) *D & C* (scrape of the lining of the womb done in an operating theatre under general anaesthetic)	☐	☐	☐

(10) Did you have to **miss work** to come to the clinic? No ☐ Yes ☐

(11) How easy is it for you to **get to the clinic** for an appointment?
☐ No problem
☐ Slightly inconvenient
☐ Inconvenient
☐ Very inconvenient

(12) Once you got to the clinic, **did you have to wait** to be seen by the gynae doctor? No ☐ Yes ☐

If yes, **how long** did you have to wait? __ __ minutes

(13) Do you have to come back for **another appointment**? No ☐ Yes ☐

→ **If yes, when:** ☐☐/☐☐/☐☐☐☐ (dd/mm/yyyy) **and what for?:** _____

(14) **Today**, have you been given a contact number and/or **clinic leaflet** to take away? No ☐ Yes ☐
*(**other than** our research and consent form)*

(15) Having seen the doctor, how do you **feel now** about your symptoms? *(Please tick one answer for **each** part)*

	Not at all	Not very	Quite	Very	Extremely
a *Reassured*	☐	☐	☐	☐	☐
b *Hopeful*	☐	☐	☐	☐	☐
c *Able to cope*	☐	☐	☐	☐	☐
d *Anxious*	☐	☐	☐	☐	☐
d *Other optional*	☐	☐	☐	☐	☐

please specify: _____

(16) Did you **fill in** this questionnaire at the clinic or at home?
☐ Clinic
☐ Home

(17) We would like to **contact you** by telephone in a few weeks' time to ask **how you are feeling then** about this consultation for menstrual bleeding and any evaluations you might have had. This will be important for our comparisons between the methods of evaluation.

Would that be **acceptable to you**? No ☐ Yes ☐

If yes, could you give us a **number** where you could be contacted?
Daytime: _____ *(Home/Work)*
Evening: _____ *(Home/Work)*

If 'no', we will send you a brief questionnaire by post.

Thank you very much for your help!

• 22/2/99 **Date:** |_|_|/|_|_|/|_|_|_|_| (dd/mm/yyyy) **Trial No:** |_|_|_|_|

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

Date of birth: |_|_|/|_|_|/|_|_|_|_| (dd/mm/yyyy)

ABOUT YOUR BIOPSY

1. For this biopsy, **what explanations were you given** about the biopsy procedure by the doctor or nurse?
   - ☐ None
   - ☐ A little
   - ☐ Quite good
   - ☐ Very detailed

2. Were you given the opportunity to **ask questions** about the procedure before the biopsy? No ☐ Yes ☐

3. Were you told **when/how** you would hear the **results** of the biopsy? No ☐ Yes ☐ Don't know ☐

   if yes

   when?: in _____ (weeks), **how?** (e.g. letter, next visit, from GP):_____

4. **How long** did the biopsy take altogether, from going into the consulting room to leaving it?
   |_|_| minutes

5. Some women in our research study are having two biopsy samples taken at the same time, but by different methods. Did you have **biopsy by two methods**? No ☐ Yes ☐ Don't know ☐

6. If you had biopsy by two methods, **can you compare them**?

1st biopsy **much better** than 2nd	1st biopsy **better** than 2nd	**No difference** between 1st and 2nd	1st biopsy **worse** than 2nd	1st biopsy **much worse** than 2nd
☐	☐	☐	☐	☐

7. Did you suffer **any cramps or discomfort** with the procedure?

Not at all	Minimal	A little	A lot	Severe
☐	☐	☐	☐	☐

8. Some women find the biopsy procedure unpleasant. Overall, **how would you describe the biopsy** you had today?
   - ☐ Fine
   - ☐ Slightly unpleasant
   - ☐ Quite unpleasant
   - ☐ Very unpleasant
   - ☐ Extremely unpleasant

9. Were there any **students present** at your biopsy? No ☐ Yes ☐ → If yes, **how many**? |_|_|

10. Did you **fill in** this questionnaire at the clinic or at home? ☐ Clinic ☐ Home

11. Do you have any **other comments** about your experience of this procedure?

_____

Thank you very much for your help!

• 22/2/99 **Date:** ☐☐/☐☐/☐☐☐☐ (dd mm yyyy) **Trial No:** ☐☐☐☐

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

Date of birth: ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

ABOUT YOUR HYSTEROSCOPY

1. If you have **had a hysteroscopy** (viewing the inside of the womb through a lens on a monitor) **before today**, was that hysteroscopy done... *(If this question applies to you, please answer a and b)*

	Yes	No
a ...under general anaesthetic?	☐	☐
b ...as an outpatient?	☐	☐

2. For this hysteroscopy today, **what explanations were you given** about the hysteroscopy procedure by the doctor or nurse?
   - ☐ None
   - ☐ A little
   - ☐ Quite good
   - ☐ Very detailed

3. Did you **view the hysteroscope image** of your womb lining on the monitor screen?
   - ☐ Yes → If **yes**, which of these adjectives **describes your feelings** about the experience of viewing your own womb lining: *(Please tick only one box)*
     - ☐ Interesting
     - ☐ Off-putting
     - ☐ Worthwhile
     - ☐ Other (please specify): _____
   - ☐ No, was not offered the opportunity
   - ☐ No, did not want to watch

4. Some women find the hysteroscopy procedure unpleasant. Overall, **how would you describe the hysteroscopy** you had today (not counting the last bit, which you have already described under 'biopsy')?
   - ☐ Fine
   - ☐ Slightly unpleasant
   - ☐ Quite unpleasant
   - ☐ Very unpleasant
   - ☐ Extremely unpleasant

5. Did you **fill in** this questionnaire at the clinic or at home?
   - ☐ Clinic
   - ☐ Home

6. Do you have any **other comments** about your experience of this procedure?

_____

_____

Thank you very much for your help!

• 22/2/99 **Date:** |_|_|/|_|_|/|_|_|_|_| (dd/mm/yyyy) **Trial No:** |_|_|_|_|

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

Date of birth: |_|_|/|_|_|/|_|_|_|_| (dd/mm/yyyy)

ABOUT YOUR ULTRASOUND

(1) For this ultrasound, **what explanations were you given** about the ultrasound procedure by the doctor or nurse?

- ☐ None
- ☐ A little
- ☐ Quite good
- ☐ Very detailed

(2) Did you have the ultrasound **at the time given** on your ultrasound appointment?

- ☐ Yes
- ☐ No → If no, **how long after** the ultrasound appointment time did you actually get taken for your ultrasound? |_|_| minutes

(3) Will/did you go **straight back** to see the gynae doctor **after the ultrasound**? No ☐ Yes ☐

(4) (a) Were you given any idea of the **results** of the investigation **by the doctor** carrying out the ultrasound? No ☐ Yes ☐

→ (b) if no: were you told **when/how** you would hear the results of the ultrasound?

when?: in_____ (weeks), **how?** (e.g. letter, next visit, from GP):_____

(5) Some women find the ultrasound procedure unpleasant. Overall, **how would you describe the ultrasound** you had today?

- ☐ Fine
- ☐ Slightly unpleasant
- ☐ Quite unpleasant
- ☐ Very unpleasant
- ☐ Extremely unpleasant

(6) Did you **fill in** this questionnaire at the clinic or at home?

- ☐ Clinic
- ☐ Home

(7) Do you have any **other comments** about your experience of this procedure?

_____

_____

Thank you very much for your help!

• 26/4/99 **Date of Birth:** |__|__|/|__|__|/|__|__|__|__| (dd mm yyyy) **Trial No:** |__|__|__|__|

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

REVIEW OF CLINIC ATTENDANCE

This questionnaire is for you to fill in the day after your clinic visit that completes all the investigations you have been assigned. Please answer the questions below on |__|__|/|__|__|/|__|__|__|__|*, or the day after ......, to let us know how you have found your clinic attendance(s).*

Please write in: **Date questionnaire actually completed:** |__|__|/|__|__|/|__|__|__|__| (dd mm yyyy)

(1) Now that you have had time to think about it, how do you feel about your clinic attendance(s) regarding this bleeding problem?

How well do the following statements describe **your feelings** about the clinic?
(Please tick one box in each line)

	Not at all true	Not very true	Fairly true	Very true	Not applicable
I am more worried than before the clinic attendance	☐	☐	☐	☐	
I am glad I had the investigation	☐	☐	☐	☐	☐
I do not really understand what the doctor told me about my bleeding	☐	☐	☐	☐	
I wish I had not bothered	☐	☐	☐	☐	
I feel reassured by the visit	☐	☐	☐	☐	
I would have liked more investigation of my bleeding problem	☐	☐	☐	☐	

(2) Now that you have had your clinic visit and tests, **looking to the future, how do you feel about your health**? *Please tick one box in each line* to say how you feel **now** about your bleeding problem.

	Not at all	Not very	Quite	Very	Extremely
Fed up	☐	☐	☐	☐	☐
Hopeful	☐	☐	☐	☐	☐
Worried	☐	☐	☐	☐	☐
Angry	☐	☐	☐	☐	☐
Able to cope	☐	☐	☐	☐	☐
Disappointed	☐	☐	☐	☐	☐
Relieved	☐	☐	☐	☐	☐

PTO

(3) What aspects of your clinic visit did you find most helpful?

_____

_____

_____

(4) After any of your clinic investigations did you suffer any cramps, bleeding or discomfort after you went home from the clinic?

	None at all	Hardly any	Some	A lot	Severe
Cramps	☐	☐	☐	☐	☐
Bleeding	☐	☐	☐	☐	☐
Discomfort	☐	☐	☐	☐	☐
*	☐	☐	☐	☐	☐

Any other effect (please specify)

(5) Now that you look back, how **worthwhile** do you think your clinic visit was?

Not at all	Not very	Quite	Very	Extremely
☐	☐	☐	☐	☐

(6) Do you have any comments you would like to make about the gynaecology clinic?
Please write below.

_____

_____

NB. Your answers will be completely **confidential** and will **not** be put in your medical notes.

Thank you very much for taking the time to answer these questions. Please place this questionnaire in the stamped, addressed envelope provided and post it back to us as soon as possible.

Return to:

Dr Susan Brechin
Centre for Reproductive Biology
37 Chalmers Street
Edinburgh EH3 9EW

(Telephone: 0131 229 2575)

Appendix 5

Follow-up questionnaires

Version 3
22/09/99

Date of Birth: ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

Trial No: ☐☐☐☐

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

10 MONTH REVIEW OF CLINIC ATTENDANCE

This questionnaire is for you to fill in 10 months after we first spoke to you about the study. Please answer the questions below to let us know how you feel about your clinic attendance(s) and original symptoms.

Please write in: **Today's date:** ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

1 You came to the hospital clinic 10 months ago because of menstrual bleeding problems.
How are these symptoms now (i.e. over the last month or so)?

Worse	No different	Improved	Much improved
☐	☐	☐	☐

If 'improved/much improved':

a) How long have they been improved?: for ☐☐ *months*

b) Are you troubled with the symptoms at all now? ☐ *no* ☐ *yes* *if 'No", go to Qu.3*

2 Thinking about the symptoms that brought you to the clinic, **how do you feel now** about any of these symptoms you still have:

	Not at all	Not very	Quite	Very	Extremely
Worried	☐	☐	☐	☐	☐
Able to cope	☐	☐	☐	☐	☐
Angry	☐	☐	☐	☐	☐
Reassured	☐	☐	☐	☐	☐
Dissatisfied	☐	☐	☐	☐	☐

3 You entered the study on ____ / ____ / _____ .
Since then, **how many further appointments** have you had at the gynaecology clinic, *for the same bleeding problem*?

☐☐ *further appointments*
(Enter 0, if none)

4 When was your **most recent appointment** at the gynaecology outpatient clinic?
(*please tick one*)

- ☐ Within the last month
- ☐ 2–3 months ago
- ☐ 4–6 months ago
- ☐ More than 6 months ago

5 Do you have **another appointment** at the gynaecology clinic?

- ☐ no
- ☐ yes

6 *In the past 10 months*, have you spent any time as a **patient in a hospital ward** because of your bleeding problem?

- ☐ no
- ☐ yes

If 'yes', how many days in total? ☐☐ *days in hospital*

And how many nights in total? ☐☐ *night in hospital*
(*Enter 0, if none*)

7 *In the past 10 months*, how many times have you **seen your GP** about this problem?

☐☐ *visits to GP*
(*Enter 0, if none*)

8 *In the past 10 months*, has your GP **prescribed medication** for this problem?

- ☐ no
- ☐ yes

If 'yes', please give details:

Name of medication(s): and how long taken?
(enter a number in the boxes and whether weeks or months)

..... *taken for* ☐☐ *weeks/months (please circle one)*

..... *taken for* ☐☐ *weeks/months (please circle one)*

..... *taken for* ☐☐ *weeks/months (please circle one)*

9 *In the past 10 months*, have you **had to take time off wor**k because of your bleeding problem?

- ☐ no
- ☐ yes

If 'yes', how many days in total? ☐☐ *days off work*

10 Thinking back to your attendance(s) at the gynaecology clinic, starting 10 months ago, please say **how helpful you found the various parts of your care**:

(*Please tick one answer for each part*)

	Happened			
	Did not happen	but not at all helpful	and helpful	and very helpful
ⓐ Talk with GP	☐	☐	☐	☐
ⓑ Talk with a member of our OMEE study research team	☐	☐	☐	☐
ⓒ Talk with hospital doctor	☐	☐	☐	☐
ⓓ Internal examination	☐	☐	☐	☐
ⓔ Biopsy	☐	☐	☐	☐
ⓕ Hysteroscopy	☐	☐	☐	☐
ⓖ Ultrasound scan	☐	☐	☐	☐
ⓗ Other (e.g. D&C) *Please specify:*				
.....	☐	☐	☐	☐
.....	☐	☐	☐	☐
.....	☐	☐	☐	☐

11 If you were to need an investigation again for a bleeding problem, **how would you feel if the investigation recommended was…**

(*Please tick one answer for each part*)

	Very Pleased	Pleased	Would not mind	Unhappy	Very unhappy
ⓐ …clinical (internal) examination	☐	☐	☐	☐	☐
ⓑ …hysteroscopy (viewing inside of the womb through a lens on a monitor)	☐	☐	☐	☐	☐
ⓒ …biopsy (sample of tissue taken from inside your womb)	☐	☐	☐	☐	☐
ⓓ …ultrasound scan (scan of your womb)	☐	☐	☐	☐	☐
ⓔ …a general anaesthetic in a hospital theatre?	☐	☐	☐	☐	☐

12 How well do the following statements describe **your feelings** about your clinic attendance(s)? (*Please tick one box in each line*)

	Not at all true	Not very true	Fairly true	Very true
I am satisfied with the care that I received	☐	☐	☐	☐
My attendance at the clinic failed to cure my bleeding problem	☐	☐	☐	☐
I am glad I attended the clinic	☐	☐	☐	☐
I wish I had not bothered to go to the clinic	☐	☐	☐	☐
I feel reassured by my attendance	☐	☐	☐	☐
I would have liked more investigation of my bleeding problem	☐	☐	☐	☐

13 Now that you look back, how **worthwhile** do you think your clinic visit was?

Not at all ☐ Not very ☐ Quite ☐ Very ☐ Extremely ☐

14 Generally speaking, how would you describe your **current state of health**:

- ☐ Very good
- ☐ Average
- ☐ Poor
- ☐ Very poor

15 Any other comments:

_____

_____

NB. Your answers will be completely **confidential** and will **not** be put in your medical notes.

Thank you very much for taking the time to answer all of our questions. Please place this questionnaire in the pre-paid envelope provided and post it back to us as soon as possible.

Dr Susan Brechin
Centre for Reproductive Biology
37 Chalmers Street
Edinburgh EH3 9EW

(*Telephone: 0131 229 2575*)

04/01/01 **Date of Birth:** ☐☐/☐☐/☐☐☐☐ (dd mm yyyy) **Trial No:** ☐☐☐☐

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

2 YEAR REVIEW OF CLINIC ATTENDANCE

It is now 2 years since you agreed to take part in our research study. This questionnaire marks the final stage of your involvement with our research. Please answer the questions below to let us know how you are now, and your use of health services during the past 14 months.

Please write in: **Today's date:** ☐☐/☐☐/☐☐☐☐ (dd mm yyyy)

1 You came to the outpatient clinic at the Royal Infirmary 2 years ago because of menstrual bleeding problems. Compared to then, **how are these symptoms now** (i.e. over the last few months)?

Worse than 2 years ago ☐ No different ☐ Improved ☐ Much improved ☐

If 'improved/much improved':

a) How long have they been improved?: for ☐☐ *months*

b) Are you troubled with the symptoms at all now? ☐ *no* ☐ *yes* *if 'No", go to Qu.3*

2 Thinking about the bleeding problem that brought you to the clinic 2 years ago, **how do you feel now** about any of the symptoms you still have:

	Not at all	Not very	Quite	Very	Extremely
Worried	☐	☐	☐	☐	☐
Able to cope	☐	☐	☐	☐	☐
Angry	☐	☐	☐	☐	☐
Reassured	☐	☐	☐	☐	☐
Dissatisfied	☐	☐	☐	☐	☐

3 Generally speaking, how would you describe your **state of health now**:

☐ Very good
☐ Average
☐ Poor
☐ Very poor

4 We last contacted you in ..... . Since then (i.e. over the past 14 months) **how many further appointments** have you had at the gynaecology clinic at the **Royal Infirmary**, *for the same bleeding problem*?

☐☐ *further appointments*

(Enter 0, if none)

5 When was your **most recent appointment** at the Royal Infirmary gynaecology outpatient clinic? (*please tick one*)

- ☐ Within the last month
- ☐ 2–3 months ago
- ☐ 4–6 months ago
- ☐ More than 6 months ago

6 Do you have **another appointment** at the gynaecology clinic?

- ☐ no
- ☐ yes

7 In the past 14 months have you attended a **gynaecology outpatient clinic** at another hospital (other than the Royal Infirmary)?

- ☐ no
- ☐ yes

If 'yes', please give details:

Name of hospital:	**Date of visit:** (if you cannot remember the exact date, month and year are sufficient)
.....	_____/_____/_____
.....	_____/_____/_____
.....	_____/_____/_____

8 *In the past 14 months*, have you spent any time as a **patient in a hospital ward** because of your bleeding problem, or for treatment of it?

- ☐ no
- ☐ yes

If 'yes', how many days in total? ☐☐ *days in hospital*

And how many nights in total? ☐☐ *nights in hospital*

(*Enter 0, if none*)

9 *In the past 14 months*, how many times have you **seen your GP** about your bleeding problem?

☐☐ *visits to GP*

(*Enter 0, if none*)

10 *In the past 14 months*, has your **GP prescribed medication** for this problem?

☐ no
☐ yes

If 'yes', please give details:

Name of medication(s): and how long taken?
(enter a number in the boxes and whether weeks or months)

..... *taken for* ☐☐ *weeks/months (please circle one)*
..... *taken for* ☐☐ *weeks/months (please circle one)*
..... *taken for* ☐☐ *weeks/months (please circle one)*

11 *In the past 14 months*, have you **had to take time off work** because of your bleeding problem?

☐ no
☐ yes

If 'yes', how many days in total? ☐☐ *days off work*

12 If you were to need an investigation again for a bleeding problem, **how would you feel if the investigation recommended was…**

(*Please tick one answer for each investigation*)

	Very Pleased	Pleased	Would not mind	Unhappy	Vey unhappy
…clinical (internal) examination	☐	☐	☐	☐	☐
…hysteroscopy (viewing inside of the womb through a lens on a monitor)	☐	☐	☐	☐	☐
…biopsy (sample of tissue taken from inside your womb)	☐	☐	☐	☐	☐
…ultrasound scan (scan of your womb)	☐	☐	☐	☐	☐
…one that needed a general anaesthetic in a hospital theatre?	☐	☐	☐	☐	☐

13 How well do the following statements describe **your feelings** about your attendance(s) at the gynaecology clinic in the Royal Infirmary for your bleeding problem?

(*Please tick one box in each line*)

	Not at all true	Not very true	Fairly true	Very true
I am satisfied with the care that I received	☐	☐	☐	☐
My attendance at the clinic failed to cure my bleeding problem	☐	☐	☐	☐
I am glad I attended the clinic	☐	☐	☐	☐
I wish I had not bothered to go to the clinic	☐	☐	☐	☐
I feel reassured by my attendance	☐	☐	☐	☐
I would have liked more investigation of my bleeding problem	☐	☐	☐	☐

14 Now that you look back, how **worthwhile** do you think your clinic visit was?

Not at all	Not very	Quite	Very	Extremely
☐	☐	☐	☐	☐

15 Any other comments:

_____

_____

NB. Your answers will be completely **confidential** and will **not** be put in your medical notes.

Thank you very much for all your help with our research study. Your participation is greatly appreciated. Please place this questionnaire in the pre-paid envelope provided and post it back to us as soon as possible.

Prof Hilary Critchley
Centre for Reproductive Biology
37 Chalmers Street
Edinburgh EH3 9EW

(*Telephone: 0131 229 2275*)

Appendix 6

Case-note review data extraction form

Trial Number ☐☐☐☐

OUTPATIENT METHODS OF ENDOMETRIAL EVALUATION (OMEE)

10 MONTH / 2 YEAR CASE NOTE REVIEW

Trial Number ☐☐☐☐ **DOB** ___/___/_____ **Randomised to:** ..... **Date entered trial** ___/___/_____

10 month cut-off date ___/___/_____ **2 years cut-off date** ___/___/_____ **Today's date** ___/___/_____

1. Date of GP referral letter ___/___/_____ **2.** Written consultation notes accessed? **No** / **Yes** **3.** Date most recently attended ___/___/_____ (OP/IP/ or investigation)

4. Any subsequent GP referral letter **(only re-referrals for same problem): N /Y** – If Y, how many ☐ **Date** ___/___/_____ **Date** ___/___/_____

5. Record any existing diagnosis (*at trial entry date*) *and current treatments*

---

6. Outcome at 10 months

- still under care ◯
- still under care but DNA ≥ 1 appointment ◯
- no mention of follow up appt or discharge ◯
- referred elsewhere ◯
  date of referral ___/___/_____
- discharged from gynae ◯
  dept date of discharge ___/___/_____

if other, please specify .....

Outcome at 2 years

- still under care ◯
- still under care but DNA ≥ 1 appointment ◯
- no mention of follow up appt or discharge ◯
- referred elsewhere ◯
  date of referral ___/___/_____
- discharged from gynae ◯
  dept date of discharge ___/___/_____

if other, please specify .....

Signed: .....

Randomised to: ..... Trial Number

APPOINTMENTS WITH CLINICIANS

Event no.	Date	Diagnosis mentioned: No Yes – specify	Time to fup indicated? (weeks)	Pelvic exam? No Yes – normal/ uncertain/abnormal	Treatment prescribed? (medical or surgical) N / Y – specify	Other
xx	13/11/2000	Yes – fibroids	8/52	Yes – normal	no	
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						

Enter in trt page

Randomised to: ..... Trial Number

LETTERS

Event no.	Date	To whom: GP/patient/ pathol/referral elsewhere/ other – specify	Diagnosis mentioned: No / Yes – specify →	Is diagnosis stated as relevant to original problem N/Y/NK	Treatment prescribed? (medical or surgical) No / Yes – specify	Other
xx	13/11/2000	GP	Yes – fibroids	Y	N	
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						

Enter in trt page

Randomised to: ..... Trial Number ⌊⌋⌊⌋⌊⌋

INVESTIGATIONS

Event no.	Date	Investigation type Pipelle/Tao/US etc	OMEE/extra	OP/IP	If IP – no. days	If non-OMEE				
						Successful? Y	Successful? N	Report? Y	Report? N	Outcome-from report/notes/letter (unsuccessful/normal/uncertain/abnormal)
xx	13/11/2000	Hysteroscopy	Extra	OP		✓			✓	Uncertain
25										
26										
27										
28										
29										
30										
31										
32										
33										
34										
35										
36										

Trial Number ☐☐☐☐

TREATMENTS

If no treatments given, write 'NONE'

Event ref no.	Drug name	Dose	Duration	Compliance: drug taken? N / Y / NK	If Yes: how long taken	Date of confirmation
31	Cyklokapron	1 g qid for 4 days	6 months	Y	≥ 3 mths	13/11/2000

Surgical Treatments:				
Event ref no.	**Date**	**Treatment**	**No. days in hosp (Daycase = 1)**	**Other info:**

Additional information .....

.....

.....

.....

Appendix 7

Clinician surveys at 23 months and after report of preliminary findings (40 months)

Background

The protocol included a survey of Royal Infirmary, Edinburgh (RIE) gynaecologists, addressing their experience of and preferences regarding methods of endometrial evaluation. It was intended to do this twice, first towards the end of recruitment, and again after circulation to them of a preliminary report of study findings as to performance of the three study outpatient methods of investigation.

Methods

Before developing the questionnaire, interviews were undertaken with senior gynaecologists within the RIE. The questionnaires addressed issues raised by the clinicians at interview, as well as matters relevant to the randomised trial being undertaken.

To maximise the likelihood of frank response it was decided that the survey should be anonymous. This meant that for those who responded to both surveys it was not possible to link their responses and obtain a more powerful assessment of changes from before to after the report.

Both surveys included all consultants and specialist registrars. These comprised 30 gynaecologists for the first survey and 36 gynaecologists for the second. The response rate was 22 questionnaires (13 consultants) for the first survey (73%) and 20 (ten consultants) for the second survey (56%). The second survey took place just after the entire department had moved to a new hospital site outside the city, the second hospital department to do so, so the lower response rate is not surprising. The small number of individuals studied (and responding) means that these findings must be interpreted with care.

Although the consultant subgroup is fairly stable in composition, there is rotation of trainees, so this subgroup will certainly have changed between one survey and the next. For the second survey eight out of ten gynaecologists responded that they had also completed the earlier questionnaire. It should be noted that at the time of both surveys some of the trainees may have been at the start or end of their training period in gynaecology, which has implications for their knowledge and experience.

Given the small sample of gynaecologists surveyed and responding, only selected results of relevance to this report are presented.

Results

Familiarity with methods of investigation

Figures 17 and 18 present the clinicians' self-reported familiarity with the outpatient methods of endometrial evaluation. The response not plotted (but evident as the remainder who would make the percentage up to 100) is those who responded that they had never done that investigation. Familiarity with ultrasound did not change. Familiarity with the Pipelle sampler increased, but as this was already the standard method of investigation, this change is unlikely to be due to the study or report. Familiarity with the Tao brush decreased markedly. The second survey was undertaken some months after study recruitment (and study assigned investigations) ended, and yet some respondents had not even heard of the Tao brush. Familiarity with hysteroscopy had marginally increased since the first survey.

Anticipated performance of methods

Figures 19 and 20 present the clinicians' self-reported expectation for the various investigative methods of percentage of investigations successfully completed. Changes in expectation were minimal. By the second survey fewer clinicians were expecting over 90% completion rates for Tao, Pipelle and hysteroscopy, which was a change in the direction of the study findings that had been reported to them in the interim. Expectations for completions of ultrasound investigations had increased, which was in the direction of the study findings. It is striking

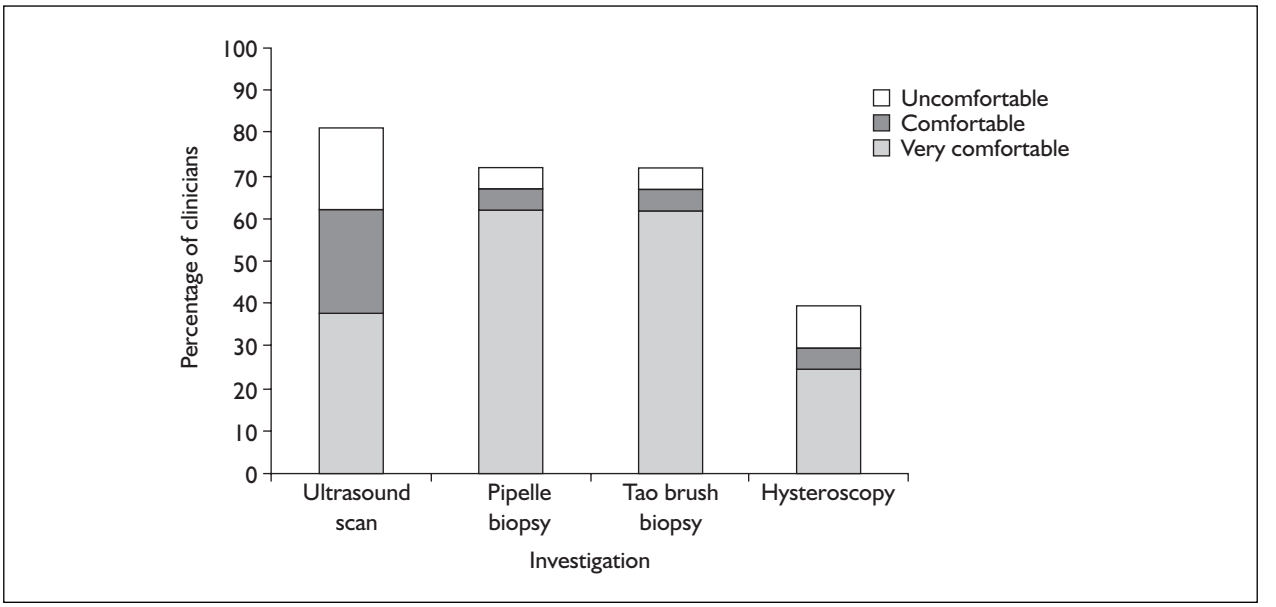


FIGURE 17 *Survey at month 23: clinician familiarity with investigations*

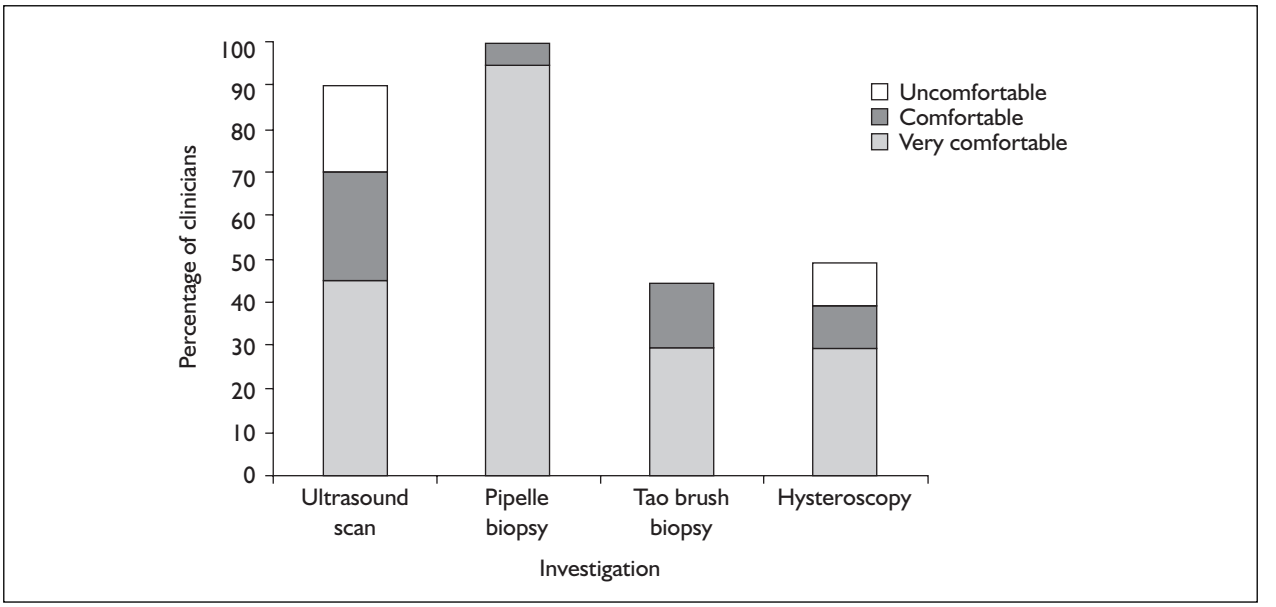


FIGURE 18 *Survey at month 40: clinician familiarity with investigations*

how wide a range of completion rates was anticipated for these investigations.

Figures 21 and 22 present the clinicians' self-reported expectation of percentage of investigations providing adequate view or, in the case of biopsy, sample material. Changes in expectation were minimal. By the second survey clinicians were expecting higher successful visualisation rates for both ultrasound and hysteroscopy, and the increase was slightly more for ultrasound, in line with study findings for relative performance for ultrasound and hysteroscopy. The expectation for adequacy of sample material for the two biopsy methods had also increased, and more so for Tao than Pipelle, in line with the study findings. However, the relative expectations for Tao and Pipelle remained more optimistic for Pipelle compared with Tao, than is the reality. Once again, it is striking how wide a range of adequacy rates was anticipated for these investigations.

Comparison of expectations reported in second survey against study findings

The study report illustrated that 84% and 74% of outpatient hysteroscopies were successfully

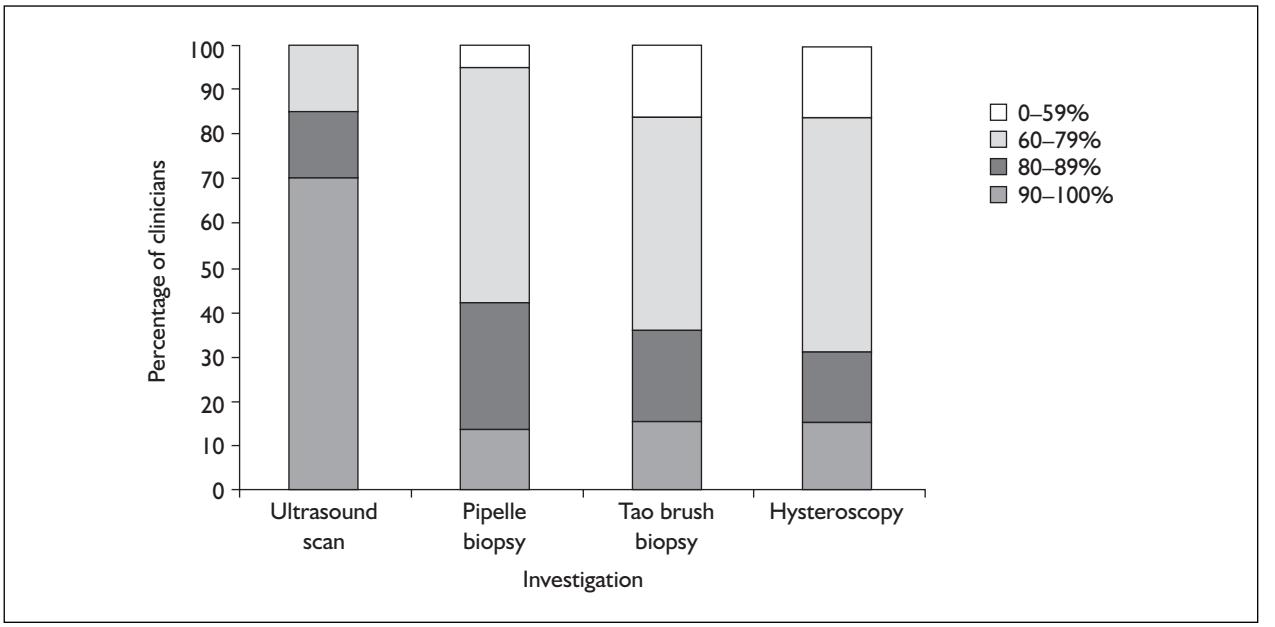


FIGURE 19 *Survey at month 23: clinician judgement of completion rate of investigation*

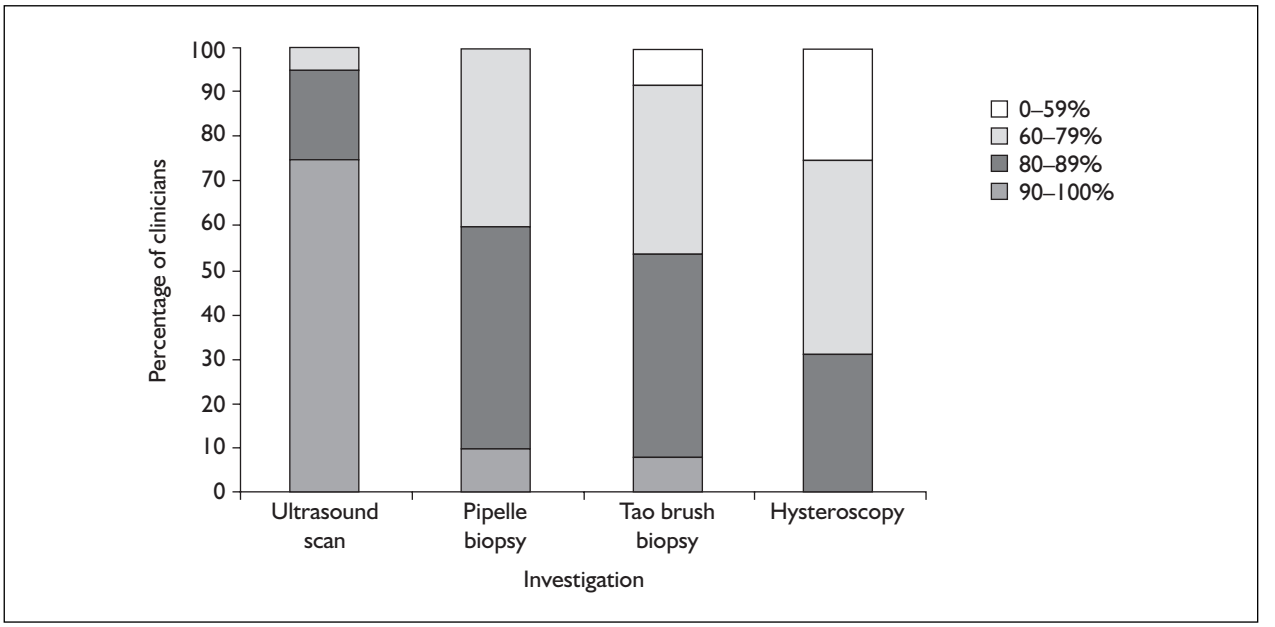


FIGURE 20 *Survey at month 40: clinician judgement of completion rate of investigation*

completed in high- and moderate-risk women, respectively. By comparing this to the relevant column of *Figure 20* it can be seen that about half of the respondents were correct in their expectation. The study figures for ultrasound were 95% and 80%, respectively, very similar to clinician responses in the survey. In the trial Pipelle samples were possible from 82% of women, so although the majority of the clinicians had a realistic view, another large subgroup underestimated performance. Similar proportions of biopsy samples were achieved for the Tao brush, and a similar interpretation of the clinicians' survey responses applies, although in this case their expectations were even less favourable than for Pipelle.

With regard to comparisons in terms of successful investigations, where they had been completed, 79% and 64% of hysteroscopies had a non-obscured view for postmenopausal women and moderate-risk women, respectively. The clinicans' expectations were considerably more optimistic than this. The study figures for adequate view at

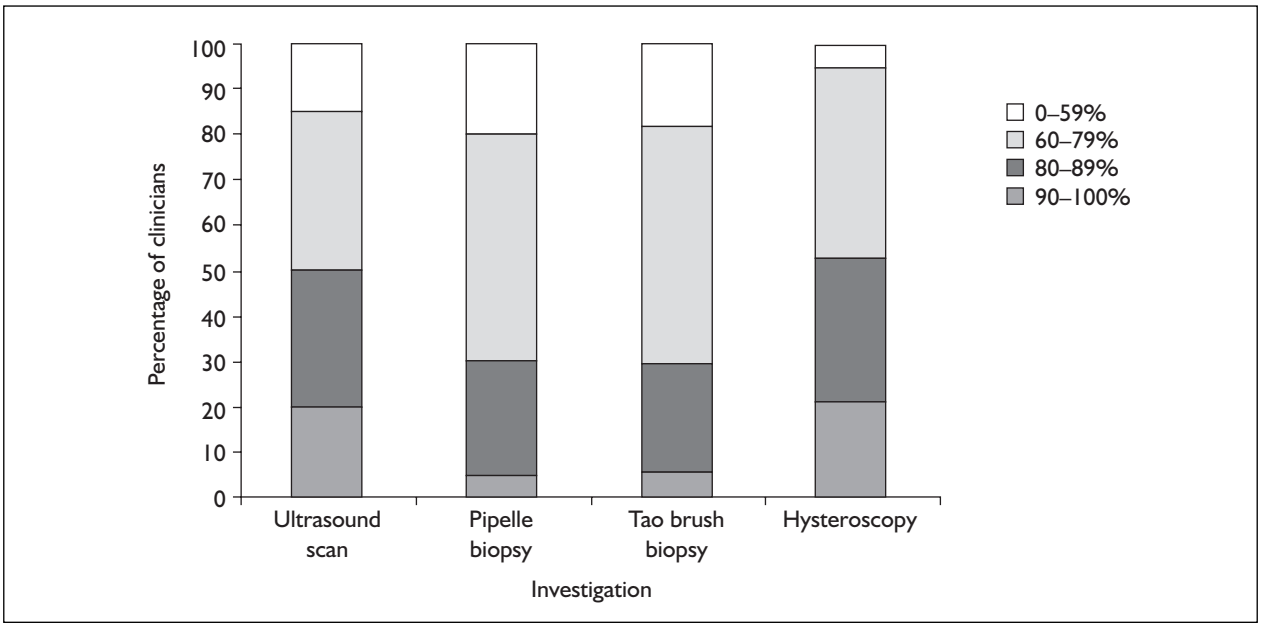


FIGURE 21 *Survey at month 23: clinician judgement of adequacy of investigation*

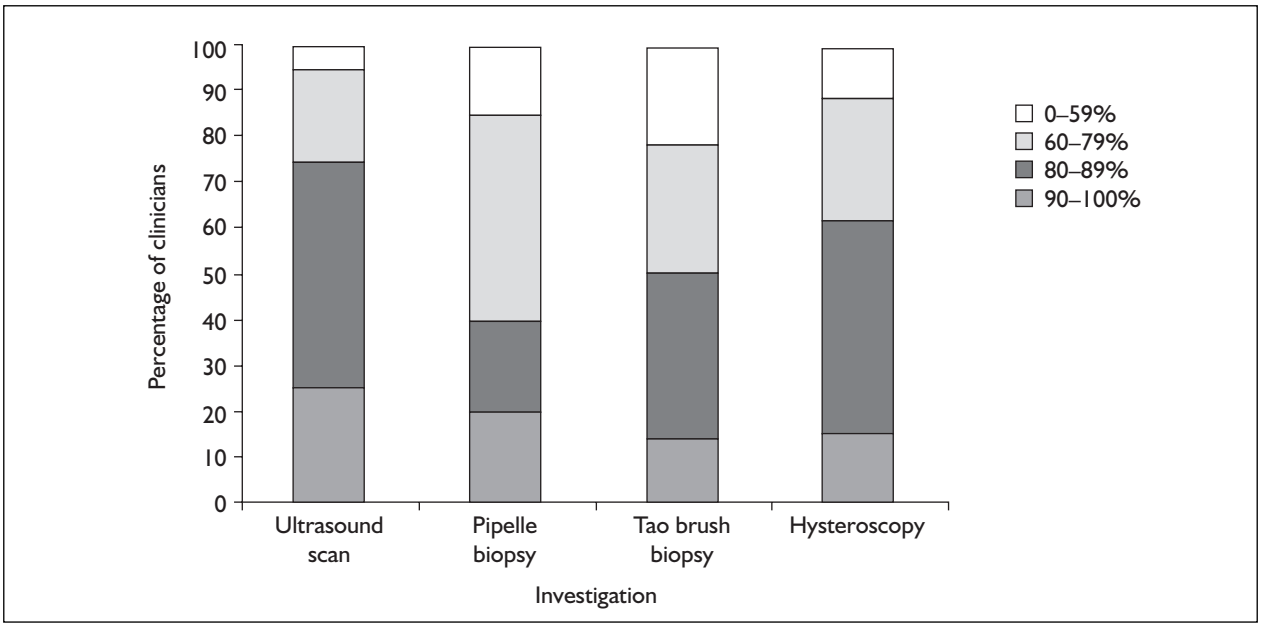


FIGURE 22 *Survey at month 40: clinician judgement of adequacy of investigation*

ultrasound were 87% and 73%, respectively, which was similar to clinicians' expectations.

With regard to successful Tao brush biopsies in the study, adequate or barely adequate samples were obtained from 88% of postmenopausal women and 86% of moderate-risk women. More than half of the clinicians underestimated this performance. For Pipelle biopsies adequate samples were obtained for 52% of completed biopsies in postmenopausal women and 91% in moderate-risk women.

If the clinicians were thinking about premenopausal women, then their expectations were very cautious, whereas if thinking of postmenopausal women, the women for whom cancer risk is greater and for whom biopsy is most used, then they were hugely optimistic.

Change in practices

In the second survey questions were added about changes to practice of investigation. Thirteen (76%) clinicians (of the 17 who answered the

TABLE 48 *Changes made within past three years since the outpatient methods of endometrial evaluation (OMEE) study started*

Greater use of ultrasound
More often depending on ultrasound alone
Pipelle and ultrasound rather than hysteroscopy as inpatient
Realisation that much of what we do in medicine is driven by technology, not by need or common sense
Use more ultrasound as primary investigation, use more hysteroscopy as backup if ultrasound results abnormal
Used ultrasound more. Refer more when indicated. Use Pipelle when indicated by ultrasound
With advent of flexible and semi-flexible fine hysteroscopes, have increased number of outpatient hysteroscopies performed as easier to perform and less discomfort

TABLE 49 *Factors that have influenced change*

Influence	**Clinician[a]**								
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Published papers and guidelines	1	5	1	3	2	2	2	1	
Your own patients randomised within this study		4	3		6		3	2	
Peer influences	2	1	2	2	5		5	6	1
Reading preliminary report for study		3			5		4	5	
Commercial advertising material		6	5		7		6	7	
Availability of services/products within gynaecological outpatient department		2	4	1	2	1	1	3	1

Factors are ranked by order of importance, with 1 = most important.
[a] Letters A–I are used for individual but anonymous respondents.

question) stated that they would like the option of using the Tao brush biopsy and indicated the groups to which this applied (multiple response possible): ten (91%) indicated high-risk women, 12 (100%) moderate-risk women, and one (12.5%) low-risk women. However, none gave a response of high-risk women only, whereas the study (and hence the report) showed that the advantage in using the Tao brush was only evident in high-risk (postmenopausal) women.

Two (14%) clinicians stated they would use the Pipelle 'much more' or 'more', ten (77%) clinicians stated they would use the Tao brush more (if available), seven (50%) clinicians would use outpatient hysteroscopy more and three (21%) would use ultrasound more. Two (14%) clinicians stated they would use the Pipelle 'less' or 'much less', two (15%) stated they would use the Tao brush less (if available), one (7%) would use outpatient hysteroscopy less and none of the clinicians would use ultrasound less.

Clinicians were asked what changes they had made in the past 3 years since the study started. Six clinicians reported making specific changes (*Table 48*).

TABLE 50 *Changes that may be made by clinicians*

Greater use of Tao brush if available
Investigate less
Use of Tao brush in high-risk women with inadequate Pipelle sample
Use of Tao brush. Referral for outpatient hysteroscopy

TABLE 51 *Requested changes to outpatient clinic*

A shorter list for menstrual problems clinic
Easier access to outpatient hysteroscopy with guidelines for referral
Flexible hysteroscopes
More access to outpatient hysteroscopy in selected cases
One-stop clinic with ultrasound, hysteroscopy, etc. would be helpful
Ultrasound facilities alongside outpatient hysteroscopy that can be performed at time of assessment, gynaecologists providing/performing ultrasounds themselves in clinics
Much more readily available outpatient hysteroscopy service. It would be great to be able to see a patient for the first time then hysteroscope them that day in the gynaecological clinics and not just menstrual problem clinics

Clinicians who reported a change over the past few years were asked to rank factors that influenced their decisions. Eight clinicians ticked at least one box in the list of factors (*Table 49*). Many of the clinicians did not rank all of the features, perhaps because they did not all apply or did not have any influence. Neither past involvement of women in the study nor the study report was considered to be a strong influence on changing practice. Peer influences on practice are claimed to be strong. The availability of services and products within gynaecology outpatient clinics, published papers and guidelines and peer influences were the most influential (highest rank, i.e. lowest number). Few clinicians stated that they were thinking about changing the way in which they investigated women with problem bleeding (*Table 50*). Approximately one-third stated methods and services that they would like changed within the gynaecology outpatient clinic (*Table 51*). Eighteen (95%) of the 19 clinicians answering the question stated that they would find a clinical pathway useful.

Appendix 8

Additional material regarding cost evaluation

Costs

Detailed costs and approaches to their calculation are provided below:

- unit costs of staff time
- unit costs of tests and radiology
- speciality costs
- allocated capital equipment costs for hysteroscopy
- staff costs for day-case procedures
- investigation/procedure costs
- cost of letters written and read
- unit costing of investigations
- calculation of drug costs.

Estimation of the unit costs of staff time

Data were taken from Personal Social Services Research Unit (PSSRU), 2001.[60] Unit costs of staff are calculated to reflect patient contact (*Table 52*).

All staff costs incorporate salary at midpoint of the scale, salary on-costs, qualifications, overheads, capital overheads (based on new build and land requirements for NHS facilities and adjusted to reflect shared use of office space for administrative, recreational and changing facilities) excluding treatment space, working time and ratio of direct to indirect time on face-to-face contacts.

Unit costs of tests and radiology

Scottish average unit costs for tests and radiology were taken from published data (*Table 53*).[61]

Specialty costs

Although average gynaecology specialty costs are available,[61] figures for day-case and inpatient include theatre, laboratory and pharmacy costs. Base costs were needed, to which the various combinations of procedures could be added. Costs used in the study are given in *Table 54*.

TABLE 53 *Unit costs of tests and radiology*

Type of test	Unit cost
Microbiology	£8.59/specimen
Clinical chemistry	£0.89/test
Haematology	£4.74/specimen
Pathology	£21.80/histopathology block
Computed tomography scan	£92.15/patient
Gamma camera	£165.97/patient
Magnetic Resonance Imaging	£107.28/patient
Ultrasound	£42.53/patient
Other radiology	£43.42/patient

TABLE 52 *Estimation of unit costs of staff time*

Grade of staff	Cost	Unit of cost
GP	£19	9.36-minute consultation
GP general medical services	£77	Hour, e.g. letter writing
Nurse grade E (staff nurse, day ward)	£32	Hour of patient contact
Nurse grade B (hospital auxiliary nurse)	£19	Hour of patient contact
CSW	£19	Hour of patient contact
ODP	£19	Hour of patient contact
Consultant pathologist	£114	Patient-related hour
Hospital surgical consultant	£114	Patient-related hour
	£199	Hour spent operating
Hospital medical consultant	£114	Patient-related hour
Consultant anaesthetist	£199	Hour spent in theatre
Hospital senior house officers	£37	Hour worked
Specialist registrar	£43	Hour worked

CSW, clinical support worker; ODP, operating department practitioner.

TABLE 54 *Specialty costs*

Type of patient	Average speciality cost	Theatre costs	Laboratory costs	Base cost
Outpatient	£57/case	–	–	£57/case
Day case	£436/case	£116	£26	£294/case
Inpatient	£1182/case/3 days	£251	£97	£278/case/day

TABLE 55 *Allocated capital equipment costs for hysteroscopy*

Type of hysteroscopy	Allocated capital equipment cost from supplier A or B
Outpatient	£52.39 (A) or £51.75 (B)/patient
Day case with autoclavable sheaths	£187.95 (A)/patient
Day case with disposable sheaths	£180.65 (A) or £184.90 (B)/patient, excluding sheaths

Information collected during the trial provided details of theatre procedures, allocated theatre costs, investigations and tests carried out; these were individually costed to permit later addition to the gynaecology speciality base costs. It was decided not to deduct the pharmacy costs from the base costs as these represent drugs supplied in the hospital setting; information from case notes provided further details of medicines prescribed for the women to take on an ongoing basis and were therefore added to the total patient costs.

Allocated capital equipment costs for hysteroscopy

The capital equipment costs for hysteroscopy are derived from the data given in *Table 55*. The UK reference cost for 2001 was £134 per patient for outpatient hysteroscopy. The central sterile supplies department (CSSD) reprocessing costs have a factor built in to allow for maintenance and replacement; however, the capital outlay associated with purchase of the major items such as hysteroscopes, couch, monitor, stacking system, camera, coupling, pulse oximeter and gas insufflator was annuitised to give an equivalent annual cost, E. Calculation of equivalent annual cost was according to Drummond and colleagues[51] and was based on the annuity factor, A; expected lifespan of each piece of equipment, n; the discount rate of 6% per annum, r; the assumption that the resale value was equal to zero, $S = 0$; and the initial purchase price, K (from the suppliers):

$$E = (K–(S/(1+r)^n)/A(n,r)$$

Thus $E = K/7.3601$, where $n = 10$.

Two suppliers, A and B, were asked for quotes for the price of equipment needed to perform hysteroscopy, one of which (A) provided quotes for a hysteroscope with a reusable sheath and with a disposable sheath, and the other (B) quoted for use with a disposable sheath only. During the study, the type of hysteroscope used was A with reusable sheath; however, it is envisaged that a new service would be likely to use type B, so castings were carried out for all cases to help to inform planning.

VAT at 17.5% was added to the quoted prices.

Calculations were based on the requirement of five sets of hysteroscopy equipment for outpatient use and six sets of equipment for day-case use, to run a service at the level experienced in the study [i.e. 205 outpatient department (OPD) patients per year and 58 day-case patients per year]; this was to allow for sterilising equipment following each use.

Capital equipment used for hysteroscopy included hysteroscope, camera head and 27-mm lens, fibre telescope, sheath (not included for day cases where a disposable sheath was used), light cable and tray; electric couch; monitors, light source, camera coupling and stacking system; D&C tray (for day cases); hysteroscopy tray (for outpatients); pulse oximeter (not included for day cases as in theatre); and gas insufflator. Factors taken into account when calculating the allocated capital equipment costs included the acquisition cost, lifespan, number of patients per year, annuitised cost and the reprocessing and maintenance costs. Hysteroscopy capital equipment costs were allocated to each patient as per *Table 56*.

TABLE 56 *Allocation of hysteroscopy capital equipment costs to each patient*

	Hysteroscope + 27-mm lens, 2-mm fibre telescope, tray, sheath (not included for day case if disposable sheath used)	**Electric couch**	**Monitors, light source, camera, controller and head, coupling, light cable stacking system**	**D&C tray (used for hysteroscopy in day case)**	**Hysteroscopy tray (only used in OPD)**	**Pulse oximeter**	**Gas insufflator**
Acquisition cost	(A): £4401.55 5: £22,008 6: £26,409 Excluding sheath 6: 24,636 (B): 5: £21,003 6: £25,204	£1600	(A): £14,602 (B): £15,275	£300	£13	£2350	£2621
Lifespan (years)	5 years	10 years	8 years	15 years	5 years	10 years	10 years
Number of patients/year	OPD: 205 DC: 58	OPD: 205 DC: 58	OPD: 205 DC: 58	OPD: 205 DC: 58	OPD: 205 DC: 58	OPD: 205 DC: 58	OPD: 205 DC: 58
Annuitised cost[a]	(A) 5: £5224.58/year (A) 6: £6269.35/year Excluding sheath (A) 6: £5846/year (B) 5: £4985.99/year (C) 6: £5983.29/year	£217/year	(A) £2351/year (B) £2460/year	£31/year	£3/year	£319.29/year	£356/year
Reprocessing CSSD/ maintenance[b]	£6.00/patient	£50/year	£100/year	£19.52/patient	£4.00/patient	£20/year	£50/year
Cost/patient OPD	(A) **£31.49** (B) **£30.32**	**£1.30**	(A) **£11.96** (B) **£12.49**	NA	**£4**	**£1.66**	**£1.98**
Cost/patient day case[c]	(A) **£114.09** Excluding sheath **£106.79** (B) **£109.16**	**£4.60**	(A) **£42.26** (B) **£44.14**	**£20**	NA	NA	£7

[a] Calculation of annuitised capital cost (*E*) is described in the text.
[b] Reprocessing CSSD/maintenance costs include the cost of sterilising, maintenance and replacement, so once the equipment is purchased initially the capital outlay will never have to be met again. Therefore, the capital outlay can be spread over an infinite period; realistically, this has been assumed to be 15 years, following which technology will probably have advanced to such an extent that effectively restarting the service will be necessary.
[c] The costs of a pulse oximeter and an anaesthetic machine were not included in day-case costing as they were assumed to be standard operating theatre equipment.

Calculation of staff costs for day-case procedures

The calculations of staff costs for day-case hysteroscopy procedures are given in *Table 57*.

TABLE 57 *Calculation of staff costs for day-case procedures*

(a) Hysteroscopy in theatre under general anaesthetic (day case)

Staff	Time (preop, procedure, recovery) per hysteroscopy	Rate	Staff cost per hysteroscopy
One grade E nurse	40 minutes	£32/hour patient contact	£21.33
Two grade D nurses[a]	1 × 40 minutes	£29/hour patient contact	£19.22
	1 × 30 minutes		£14.50
Two CSW	2 × 30 minutes	£19/hour patient contact	£19.00
One surgeon	20 minutes	£199/hour operating	£66.33
One anaesthetist	30 minutes	£199/hour theatre	£99.50
One ODP	30 minutes	£19/hour patient contact	£9.50
Total			£249.38

[a] Grade D nurse cost per hour patient contact rate calculated as the PSSRU grade E nurse rate less 9.9% as the midpoint of a grade D nurse salary is 9.9% less than that of a grade E nurse salary (the corresponding rate for a grade E nurse is published by PSSRU as £32/hour patient contact). Time for which the operating theatre is in use for each hysteroscopy was estimated at 30 minutes.

(b) D&C in theatre under general anaesthetic (day case)

Staff	Time (preop, procedure, recovery not on ward) per D&C	Rate (PSSRU + ISD)	Staff cost per D&C
One grade E nurse	40 minutes	£32/hour patient contact	£21.33
Two grade D nurses	1 × 40 minutes	£29/hour patient contact	£19.22
	1 × 30 minutes		£14.50
Two CSW	2 × 30 minutes	£19/hour patient contact	£19.00
One surgeon	20 minutes	£199/hour operating	£66.33
One anaesthetist	30 minutes	£199/hour theatre	£99.50
One ODP	30 minutes	£19/hour patient contact	£9.50
Total			£249.38

Theatre in use for 30 minutes.
ISD, Information and Statistics Division.

(c) Diagnostic laparoscopy in theatre under general anaesthetic (day case)

Staff	Time per D&C (preop, procedure, recovery not on ward)	Rate (PSSRU + ISD)	Staff cost per D&C
Two grade E nurses	1 × 60 minutes	£32/hour patient contact	£32.00
	1 × 90 minutes		£48.00
Two grade D nurses	1 × 60 minutes	£29/hour patient contact	£29.00
	1 × 90 minutes		£43.50
Two CSW	2 × 60 minutes	£19/ hour patient contact	£38.00
One surgeon	60 minutes	£199/hour operating	£199.00
One anaesthetist	90 minutes	£199/hour theatre	£298.50
One ODP	90 minutes	£19/ hour patient contact	£28.50
Total			£716.50

Investigation/procedure costs

For each type of investigation or procedure that was costed for this study (*Table 58*), the time that each grade of staff was required to devote to the conduct of the process was estimated by clinicians providing the service for the study; capital equipment costs were taken from *Table 55*; the cost of disposables was calculated from data collected from clinicians providing the service, information from suppliers and the hospital supplies department; and the costs of processing, interpreting and reporting laboratory test results were calculated from information collected from the various laboratory and pathology departments. Full costing details are shown in *Table 59*.

TABLE 58 *Investigation/procedure costs*

(a) Outpatient procedures (excluding OPD cost)

Investigation/procedure	**Cost per person (using A with reusable sheaths/B with disposable sheaths)**
Hysteroscopy (excluding biopsy)	£136.93/£194.54
Tao Brush biopsy	£93.99
Pipelle biopsy	£64.68
Tao + Pipelle biopsies	£158.68
Hysteroscopy + Pipelle biopsy	£190.53/£248.14
Hysteroscopy + Tao biopsy	£219.84/£277.45
Hysteroscopy + Pipelle + Tao biopsies	£273.44/£331.05

(b) Day-case procedures (including average gynaecology day-case cost)

Investigation/procedure	**Cost per person using reusable sheath (A)**	**Cost per person using disposable sheath (A/B)**
Hysteroscopy (excluding biopsy)	£848.97	£899.92/£904.17
Hysteroscopy + Pipelle biopsy	£902.57	£953.52/£957.77
Hysteroscopy + Tao biopsy	£931.88	£982.83/£987.08
Hysteroscopy + Pipelle + Tao biopsies	£985.48	£1036.43/£1040.68
D&C	£647.77	
Laparoscopy	£1218.38	
Cervical or vault biopsy	Normally done alongside other investigations: add pathology costs of £50.50 to these	

(b) Inpatient procedures (including average gynaecology day-case cost)

Inpatient investigations (including day 1 of inpatient stay, other days of stay to be added at £278/day)	**Cost (using A with reusable hysteroscopy sheath/using A with disposable hysteroscopy sheath/using B with disposable hysteroscopy sheath)**
Hysteroscopy	£832.97/£883.92/£888.17
Hysteroscopy + Pipelle biopsy	£886.57/£937.52/£941.77
Hysteroscopy + Tao biopsy	£915.88/£966.83/£971.08
Hysteroscopy + Pipelle + Tao biopsies	£969.48/£1020.46/£1024.68
Hysteroscopy + D&C	£1160.19/£1211.14/£1215.39
Hysteroscopy + Pipelle + Laparoscopy	£1179.70/£1230.65/£1234.90
Hysteroscopy + D&C + Laparoscopy	£1453.32/£1504.27/£1508.52

TABLE 59 *Total costs for each study investigation, undertaken as outpatient or day case*

Investigation/ procedure	Time for procedure (minutes)	Cost of medical staff at hospital surgical consultant (rate of £114/hour)	Cost of medical staff at specialist registrar rate of £43/hour	Cost of nursing staff at grade E rate of £32/hour patient contact	Cost of nursing staff at grade B rate of £19/hour patient contact	Allocated cost of capital equipment per patient	Cost of disposables per patient[a]	Cost of processing, interpreting and reporting laboratory results, based on laboratory + pathologists time at £114/hour	Total cost per patient (range based on staff grades involved: always one doctor and one or two nurses throughout) excluding OPD appointment
Outpatient hysteroscopy (excluding biopsy)	30	£57.00	Zero to allow for solo/joint working	£16	£9.50	(A) £52.39 (B) £51.75	(A): (£1 + 54p + 50p)= £2.04 (B): (£1+54p + £58.75) = £60.29	NA	(A) £136.93 (B) £194.54
Day-case hysteroscopy costs (excluding staff)	See below	NA	NA	NA	NA	(A) £180.65 or £187.95 (including reusable sheath) (B) £184.90	(A): (£1 + £20 + £58.75) = £79.75 or (£1 + £20 + 50p) = £21.50 (B): £79.75	NA	(A) Capital equipment + disposable = £209.45 including reusable sheath (+ Cidex) or (A) £260.40 (B) £264.65 including disposable sheath and excluding reusable sheath See below
Day-case hysteroscopy costs (excluding staff)	See below	NA	NA	NA	NA	(A) £180.65 or £187.95 (including reusable sheath) (B) £184.90	(A): (£1 + £20 + £58.75) = £79.75 or (£1 + £20 + 50p) = £21.50 (B): £79.75	NA	(A) Capital equipment + disposable = £209.45 including reusable sheath (+ Cidex) or (A) £260.40 (B) £264.65 including disposable sheath and excluding reusable sheath See below
Biopsy (Tao)	5	£9.50	Zero to allow for solo/joint working	NA	£1.58	Included above	£17.63	£22 + £5.28 + £38 (20 minutes) = £65.28	£93.99
Biopsy (Pipelle)	5	£9.50	Zero to allow for solo/joint working	NA	£1.58	Included above	£3.10	£22 + £28.50 (15 minutes) = £50.50	£64.68
Tao + Pipelle	10	£19.00	Zero to allow for solo/joint working	NA	£3.17	Included above	£20.73	£115.78 (£65.28 + £50.50)	£158.68

continued

TABLE 59 *Total costs for each study investigation, undertaken as outpatient or day case (cont'd)*

Investigation/ procedure	Time for procedure (minutes)	Cost of medical staff at hospital surgical consultant (rate of £114/hour)	Cost of medical staff at specialist registrar rate of £43/hour	Cost of nursing staff at grade E rate of £32/hour patient contact	Cost of nursing staff at grade B rate of £19/hour patient contact	Allocated cost of capital equipment per patient	Cost of disposables per patient[a]	Cost of processing, interpreting and reporting laboratory results, based on laboratory + pathologists time at £114/hour	Total cost per patient (range based on staff grades involved: always one doctor and one or two nurses throughout) excluding OPD appointment
OPD hysteroscopy + Biopsy (Pipelle)	30	£57	Zero to allow for solo/joint working	£16.00	£9.50	(A) £52.39 (B) £51.75	(A): (£1 + 54p + 50p) = £2.04 + £3.10 = £5.14 (B): (£58.75 + 54p + £1) = £60.29 + £3.10 = £63.39	£50.50	(A) £190.53 (B) £248.14
Day-case hysteroscopy + Biopsy (Pipelle) (excluding staff)	See below	NA	NA	NA	NA	(A) £180.65 or £187.95 (including reusable sheath) B) £184.90	(A): £79.75 + £3.10 = £82.85 or £21.50 + £3.10 = £24.60 (B): £79.75 + £3.10= £82.85	£50.50	(A) Capital equipment + disposable+pathology = £263.05 (including reusable sheath + Cidex) or (A) £314 (B) £318.25 (including disposable sheath) See below
OPD hysteroscopy + Biopsy (Tao)	30	£57	Zero to allow for solo/joint working	£16.00	£9.50	(A) £52.39 (B) £51.75	(A): (£17.63+ £2.04) = £19.67 (B): (£1 + 54p + £58.75) = £60.29 + £17.63 = £77.92	£65.28	(A) £219.84 (B) £277.45
Day-case hysteroscopy + Biopsy (Tao) (excluding staff)	See below	NA	NA	NA	NA	(A) £180.65 or £187.95 (including reusable sheath) (B) £184.90	(A): £79.75 + £17.63 = £97.38 or £21.50+ £17.63 = £39.13 (B): £79.75+ £17.63 = £97.38	£65.28	(A) Capital equipment + disposables + pathology = £292.36 (including Cidex + reusable sheath) or (A) £343.31 (B) £347.56 (including disposable sheath) See below

continued

TABLE 59 *Total costs for each study investigation, undertaken as outpatient or day case (cont'd)*

Investigation/ procedure	Time for procedure (minutes)	Cost of medical staff at hospital surgical consultant (rate of £114/hour)	Cost of medical staff at spec. reg rate of £43/hour	Cost of nursing staff at grade E rate of £32/hour patient contact	Cost of nursing staff at grade B rate of £19/hour patient contact	Allocated cost of capital equipment per patient	Cost of disposables per patient[a]	Cost of processing, interpreting and reporting laboratory results, based on laboratory + pathologists time at £114/hour	Total cost per patient (range based on staff grades involved: always one doctor and one or two nurses throughout) excluding OPD appointment
OPD hysteroscopy + Biopsy (Pipelle) + Biopsy (Tao)	30	£57	Zero to allow for solo/joint working	£16.00	£9.50	(A) £52.39 (B) £51.75	(A): (£3.10 + £17.63 + £2.04) = £22.77 (B): (£1 + 54p + £58.75) = £60.29 + £17.63 + £3.10 = £81.02	£115.78	(A) £273.44 (B) £331.05
Day-case hysteroscopy + Biopsy (Pipelle) + Biopsy (Tao) (excluding staff)	See below	NA	NA	NA	NA	(A) £180.65 or £187.95 (including reusable sheath) (B) £184.90	(A): £79.75 + £17.63 + £3.10 = £100.48 or £21.50 + £17.63 + 3.10 = £42.23 (B): £79.75 + £17.63 + £3.10 = £100.48	£115.78	(A) Capital equipment + disposables +pathology = £345.96 (including Cidex + reusable sheath) or (A) £396.91 (B) £401.16 (including disposable sheath) See below

[a] Disposables include any or all of the following (depending on procedure): sterile gloves, £1.00/patient; disposable speculum, £0.54/patient (not used for day-case hysteroscopy; use a D&C tray instead at £20/patient); TAO brush, £17.63/patient (brush and fluid); endometrial sampler, £3.10/patient; disposable sterile hysteroscopy sheath, £58.75 each (now routinely used in all day-case work; in OMEE trial reusable sheaths were sterilised in-house, this is no longer permitted following relocation to the New RIE); costings were calculated to include this (indicating current and future practice) and to exclude it (in-house sterilisation costs of using Cidex were estimated at 50p/use).

Cost of letters written and read

For all patients the cost of letters written to and read by a range of individuals concerned with their care was included as an element in the calculation of total cost per case (*Table 60*). Time taken to read or write a letter was estimated at 3 minutes per letter and costed at published rates.[60]

TABLE 60 *Cost of letters*

Originator/recipient of letter	Cost of reading + writing letter
GP	£8.75
Patient	£5.00
Pathologist	£10.00
Referral elsewhere	£10.00
Family planning clinic	£10.00
Company	£3.85
Associated referral	£10.00

Unit costing of investigations

The costs of all investigations were calculated and the individual costs are given in *Table 61*.

TABLE 61 *Unit costing of investigations*

Investigation	Unit cost
Cervical biopsy[a]	£50.50
Ultrasound scan	£42.53
D&C[a]	£647.77
Suction curettage[a]	£647.77
Laparoscopy[a]	£1218.38
Colposcopy[a,b]	£114
Cervical biopsy[c]	£50.50
Ultrasound scan	£42.53
D&C[c]	£647.77
Suction curettage[c]	£647.77
Laparoscopy[c]	£1218.38
Colposcopy[b]	£114
Cervical smear	£8.59
Chlamydia swab	£8.59
High vaginal swab	£8.58
Thyroid function test	£0.89
Clotting screen/anticoagulation screen	£4.74
Haemoglobin	£4.74
Ferritin	£0.89
Hormone assays	£0.89
Luteinising hormone/follicle-stimulating hormone	£0.89
Prolactin	£0.89
Urine tracking	£10.53
Serum androgens/sex hormone binding globulin	£0.89
Electrocardiogram	£84.00
Blood tests/full blood count	£4.74
Hormone assays oestradiol/progesterone	£0.89
Pre-general anaesthetic blood tests	£4.74
Urine + electrolytes	£0.89
Liver function test	£0.89
Tumour marker (CA125, CEA)	£19.25 (together) CA125 £10; CEA £6
Midstream urine flow	£0.89
Chest X-ray	£13.00
Glucose only	£0.89
Rubella	£7.00
Group blood and save serum	£4.74
C-spine X-ray	£43.42
Pregnancy test	£4.74

continued

TABLE 61 *Unit costing of investigations (cont'd)*

Investigation	Unit cost
Venogram	£43.42
Yeast	£8.59
Actinomycetes	£8.59
Pain assessment	£57
Vulval biopsy	£21.80
Barium enema	£43.42
Blood cultures	£4.74
Magnetic resonance imaging	£107.28
Echocardiogram	£121.00
Emergency triage	£46.00

Sources for the following unit costs included the various hospital laboratories and departments, [a]the ISD of the Common Services Agency[61] and [b]the National Schedule of reference costs for the NHS in England and Wales.[69]
[c] Costed specifically for this study using information from a range of sources.

TABLE 62 *Calculation of drug costs*

Drug	Median cost
Prostaglandin synthetase inhibitors (non-steroidal anti-inflammatory) Mefenamic acid (Ponstan/Dysman/Meflam/Opustan) Naproxen (Naprosyn/Nycopren/Synflex/Napratec) Indomethacin (Indocil/Indomax) Ibuprofen (Brufen) Flurbiprofen (Froben) Diclofenac (Voltarol/Arthrotec/Diclomax)	£1.92/week
Progestogens (continuous/cyclical 14 days/cyclical 21 days) Norethisterone (Primolut N/Utovlan) (can be daily, 21 days, 10 days) Medroxyprogesterone (Provera) (10 days) = 15 Dydrogesterone (Duphaston) (21 days) = 16	£1.86/week
Oral contraceptives (list in British National Formulary) Depo-provera injections (Noristerat) (also used in malignant disease)	£1.84/week (including GP to give injection) £2.04/week
Antifibrinolytic agents Tranexamic acid (Cyklokapron) Ethamsylate (Dicynene)	£1.65/week
Clomiphene (Clomid/Serophene)	£1.77/week (max. 3 months)
Danazol (Danol)	£7.36/week (max. 6 months)
Gestrinone (Demetriose) Thyroid replacement treatment HRT (cyclical) HRT (combined)	£28.17/week (6 months) £0.53/week £1.54/week £2.31/week
LHRH analogue (Goserelin/Zoladex) Adjust anticoagulation dosage Topical HRT Antibiotic (Anti)fungal infection (Gynodaktarin) Ferrous sulfate HRT (type unknown) Raloxifene Dermovate (local steroid cream for lichen sclerosis) Progesterones Oestrogen-only HRT Other non-gynaecological drugs (postop, etc.)	£30.81/week £19 (for GP appointment) £3.89/week £2.44/week £1.91/week £0.72/week £1.93/week £5.02/week £2.12/week (max. 4 weeks) £39.46 one-off cost £0.99 week Including in day-case costs

Calculation of drug costs

Table 62 provides the costs used for the prescribed drugs taken by the women in the study. A cost for the dispensing fee of £0.946 per item was added to the monthly drug cost (from Ref. 70) and divided by 4 to give a weekly cost, apart from the case of the oral contraceptive pill which is normally dispensed once every 3 months; therefore, the dispensing fee was added to the 3-monthly drug cost and the total divided by 12 to give a weekly cost. The cost per patient was calculated according to use of the drug as recorded in the notes, or where missing, for the standard course (or until the end of the trial if prescribed indefinitely).

The median value of a range of treatment costs was used for each code as data were highly skewed; therefore, it was inappropriate to use the mean.

Sensitivity analyses

As described in Methods, sensitivity analyses were undertaken for total costs to last investigation if an investigation cost was varied by 10% over or under the cost assigned. *Tables 63–67* show the sensitivity limits obtained if the cost of hysteroscope, Pipelle biopsy or Tao brush biopsy were varied in this way, for high- and moderate-risk women (combined), and for low-risk women.

TABLE 63 *Cost-effectiveness sensitivity analysis I within moderate- and high-risk women: cost to last investigation where Pipelle biopsy is increased/decreased by 10%*

Randomised evaluation option	High risk		Moderate risk	
	n	£	*n*	£
B			78	481/468
H + B	99	638/625	81	692/679
U + B	97	726/713	77	608/595
U + H + B			80	644/631

H, hysteroscopy; U, TVUS; B, both biopsies (Tao and Pipelle).

TABLE 64 *Cost-effectiveness sensitivity analysis II within moderate- and high-risk women: cost to last investigation where Tao brush is increased by 10%*

Randomised evaluation option	High risk		Moderate risk	
	n	£	*n*	£
B			78	483/465
H + B	99	641/622	81	695/677
U + B	97	729/711	77	610/593
U + H + B			80	647/629

B, both biopsies (Tao and Pipelle).

TABLE 65 *Cost-effectiveness sensitivity analysis I and II within low-risk women: costs to last investigation where Pipelle biopsy and Tao brush are increased/decreased by 10%*

Randomised evaluation option	*n*	I Pipelle	II Tao brush
		Low risk (biopsy methods combined)	Low risk (biopsy methods combined)
None	59	749/746	748/748
B	32	456/448	456/448
H+B	30	834/827	835/826
U	32	379/377	378/378

B, Tao or Pipelle biopsy.

TABLE 66 *Cost-effectiveness sensitivity analysis III within moderate- and high-risk women: cost to last investigation where hysteroscopy is increased/decreased by 10%*

Randomised evaluation option	High risk		Moderate risk	
	n	£	*n*	£
B			78	476/472
H + B	99	647/617	81	701/671
U + B	97	734/716	77	604/600
U + H + B			80	652/624

B, both biopsies (Tao and Pipelle).

TABLE 67 *Cost-effectiveness sensitivity analysis III within low-risk women: cost to last investigation where hysteroscopy is increased/decreased by 10%*

Randomised evaluation option	*n*	Low risk (biopsy methods combined)
None	59	752/744
B	32	453/450
H+B	30	845/815
U	32	380/377

B, Tao or Pipelle biopsy.

Appendix 9

Flowcharts of progress through the trial

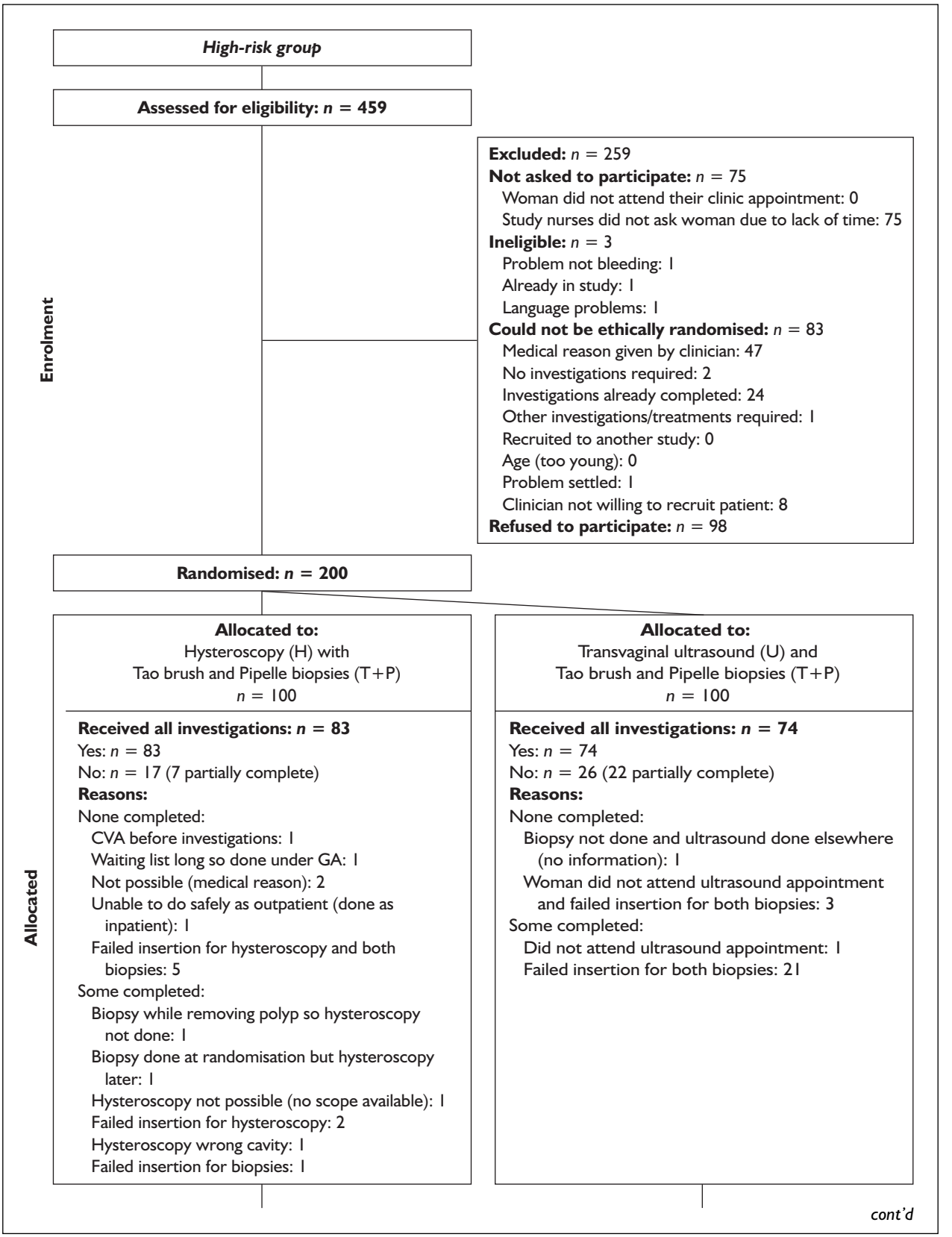


FIGURE 23 *High-risk group (postmenopausal): flowchart of progress through the trial.* [a]*Where case-note review was not 'completed' this was because patient case notes could not be located. CVA, cerebrovascular accident; GA, general anaesthetic*

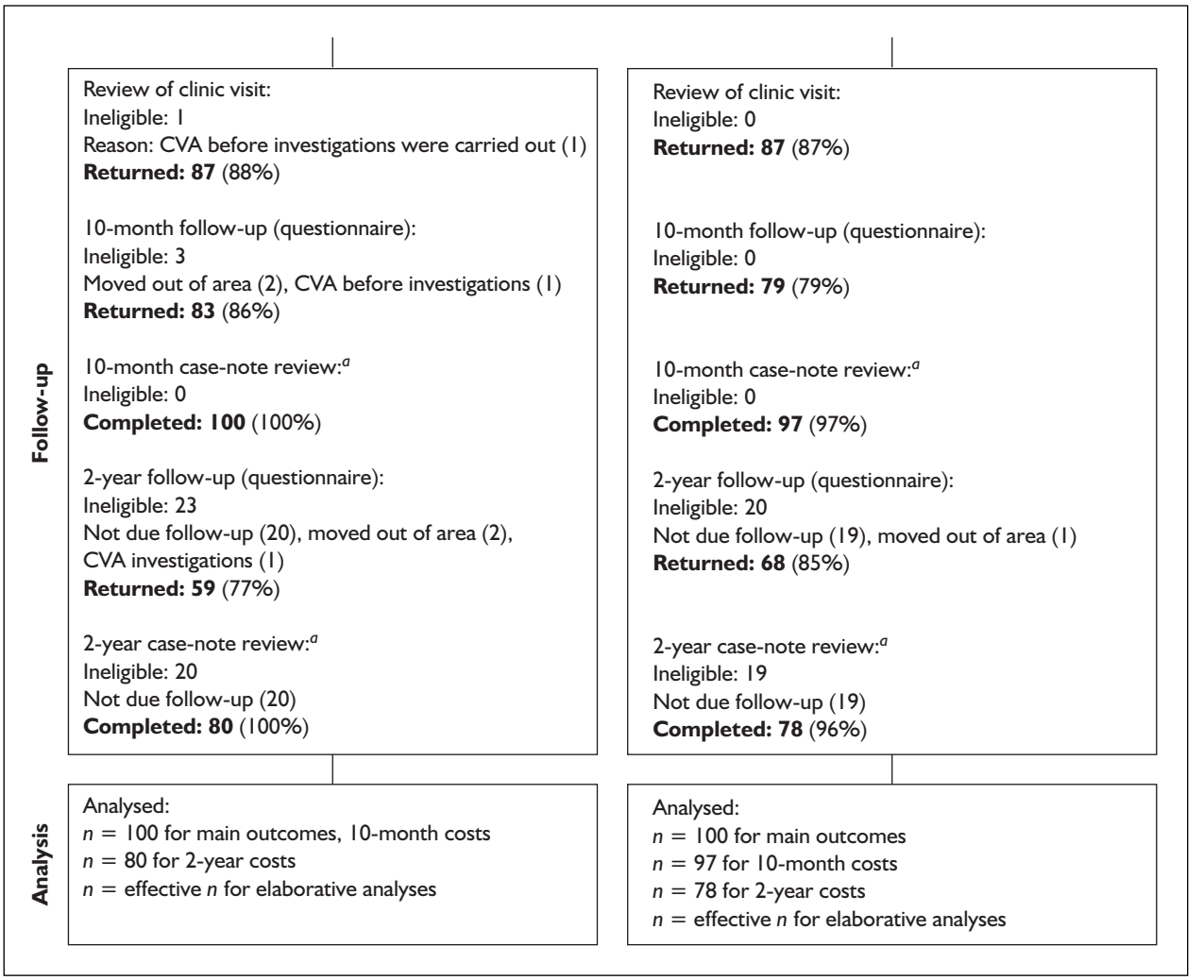


FIGURE 23 (cont'd) *High-risk group (postmenopausal): flowchart of progress through the trial.* [a]*Where case-note review was not 'completed' this was because patient case notes could not be located. CVA, cerebrovascular accident; GA, general anaesthetic*

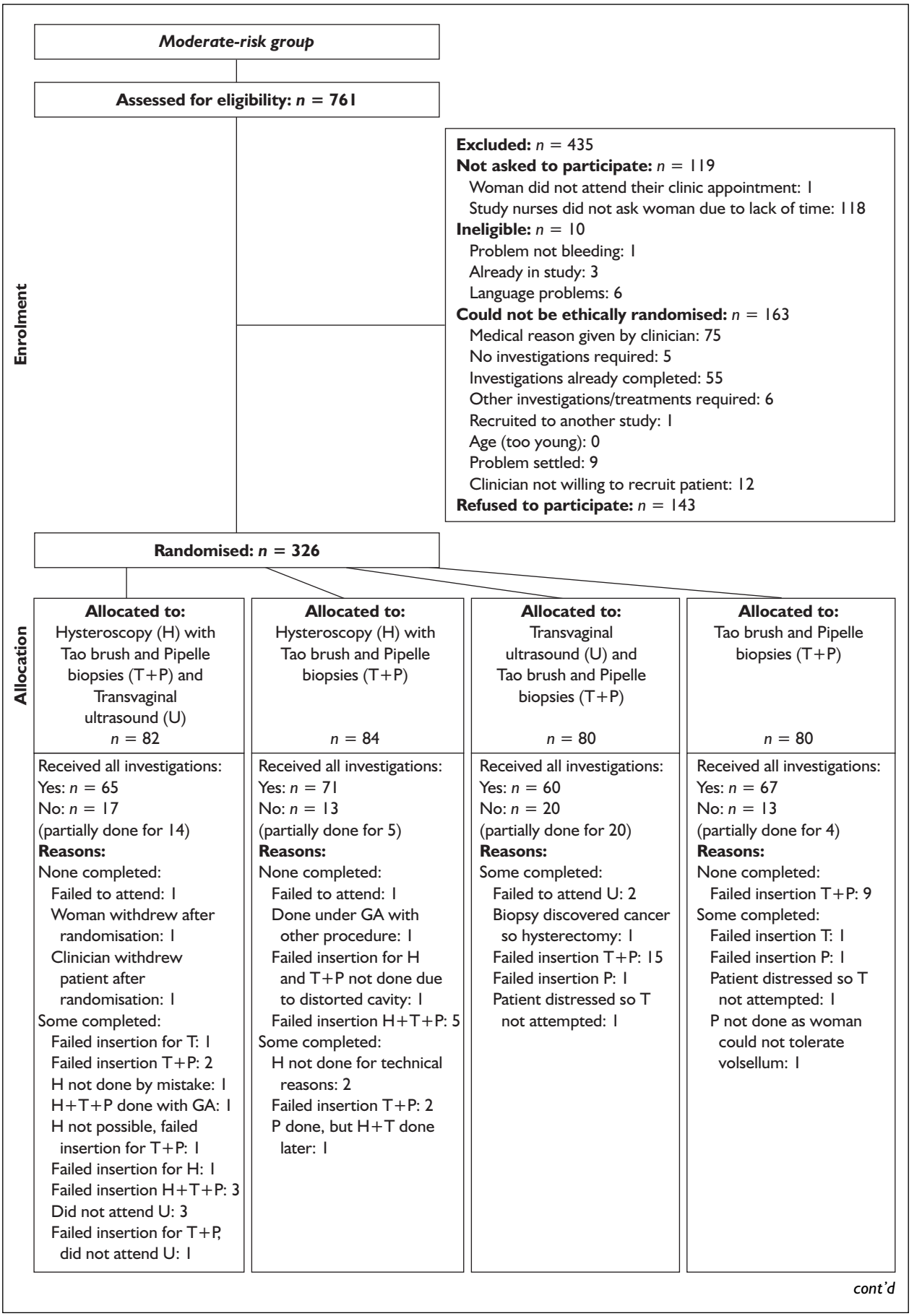


FIGURE 24 *Moderate-risk group: flowchart of progress through the trial (premenopausal and aged ≥ 40 years, or aged <40 years but with risk factors).* [a]*Where case-note review was not 'completed' this was because patient case notes could not be located*

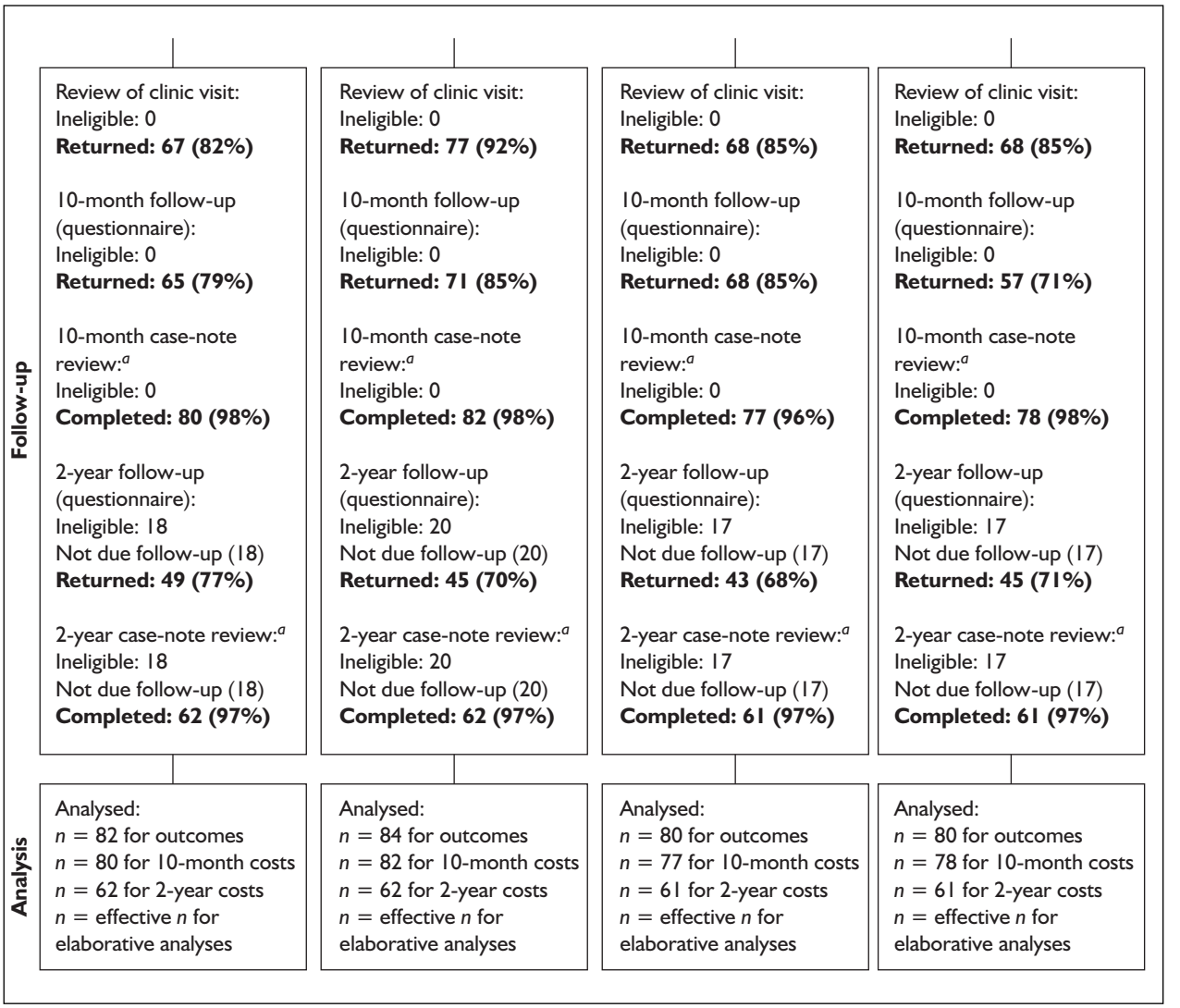


FIGURE 24 (cont'd) *Moderate-risk group: flowchart of progress through the trial (premenopausal and aged ≥ 40 years, or aged <40 years but with risk factors).* [a]*Where case-note review was not 'completed' this was because patient case notes could not be located*

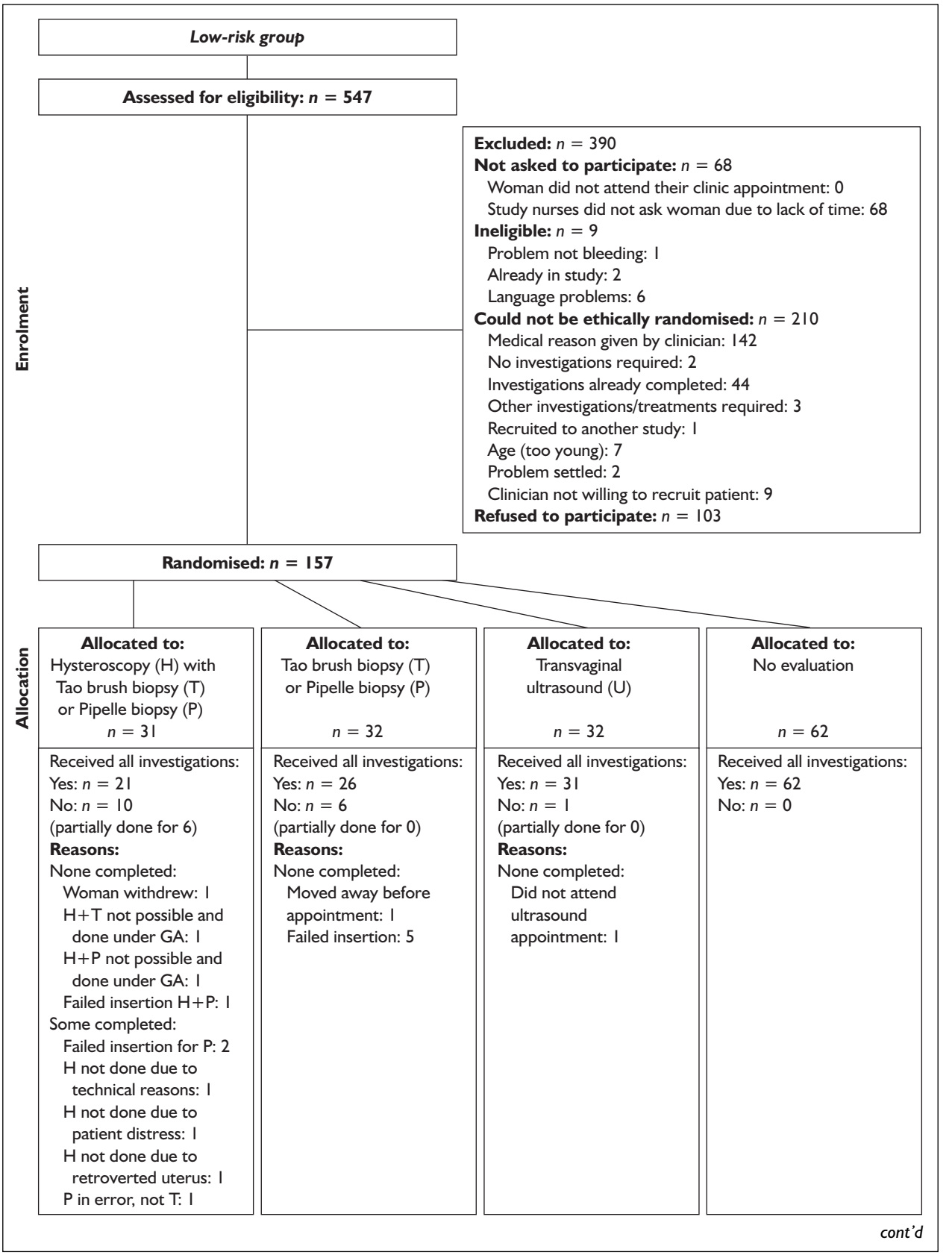


FIGURE 25 *Low-risk group: flowchart of progress through the trial (premenopausal, aged <40 years and without risk factors).*
**Where case-note review was not 'completed' this was because patient case notes could not be located*

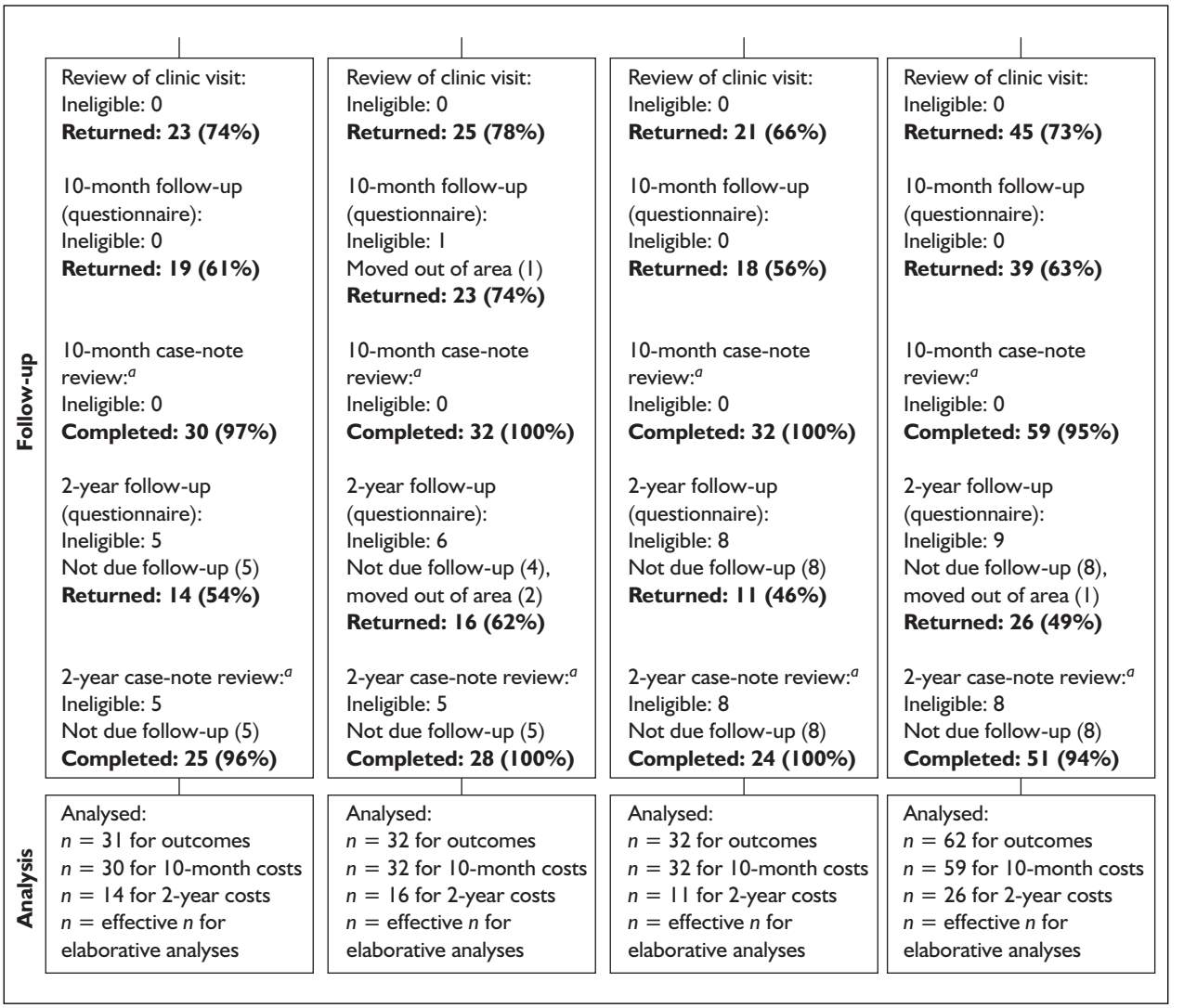


FIGURE 25 (cont'd) *Low-risk group: flowchart of progress through the trial (premenopausal, aged <40 years and without risk factors). *Where case-note review was not 'completed' this was because patient case notes could not be located*

Health Technology Assessment Programme

Prioritisation Strategy Group

Members

Chair,
Professor Tom Walley,
Director, NHS HTA Programme, Department of Pharmacology & Therapeutics, University of Liverpool

Professor Bruce Campbell, Consultant Vascular & General Surgeon, Royal Devon & Exeter Hospital

Professor Shah Ebrahim, Professor in Epidemiology of Ageing, University of Bristol

Dr John Reynolds, Clinical Director, Acute General Medicine SDU, Radcliffe Hospital, Oxford

Dr Ron Zimmern, Director, Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

HTA Commissioning Board

Members

Programme Director,
Professor Tom Walley,
Director, NHS HTA Programme, Department of Pharmacology & Therapeutics, University of Liverpool

Chair,
Professor Shah Ebrahim,
Professor in Epidemiology of Ageing, Department of Social Medicine, University of Bristol

Deputy Chair,
Professor Jenny Hewison,
Professor of Health Care Psychology, Academic Unit of Psychiatry and Behavioural Sciences, University of Leeds School of Medicine

Dr Jeffrey Aronson
Reader in Clinical Pharmacology, Department of Clinical Pharmacology, Radcliffe Infirmary, Oxford

Professor Ann Bowling, Professor of Health Services Research, Primary Care and Population Studies, University College London

Professor Andrew Bradbury, Professor of Vascular Surgery, Department of Vascular Surgery, Birmingham Heartlands Hospital

Professor John Brazier, Director of Health Economics, Sheffield Health Economics Group, School of Health & Related Research, University of Sheffield

Dr Andrew Briggs, Public Health Career Scientist, Health Economics Research Centre, University of Oxford

Professor Nicky Cullum, Director of Centre for Evidence Based Nursing, Department of Health Sciences, University of York

Dr Andrew Farmer, Senior Lecturer in General Practice, Department of Primary Health Care, University of Oxford

Professor Fiona J Gilbert, Professor of Radiology, Department of Radiology, University of Aberdeen

Professor Adrian Grant, Director, Health Services Research Unit, University of Aberdeen

Professor F D Richard Hobbs, Professor of Primary Care & General Practice, Department of Primary Care & General Practice, University of Birmingham

Professor Peter Jones, Head of Department, University Department of Psychiatry, University of Cambridge

Professor Sallie Lamb, Research Professor in Physiotherapy/Co-Director, Interdisciplinary Research Centre in Health, Coventry University

Professor Julian Little, Professor of Epidemiology, Department of Medicine and Therapeutics, University of Aberdeen

Professor Stuart Logan, Director of Health & Social Care Research, The Peninsula Medical School, Universities of Exeter & Plymouth

Professor Tim Peters, Professor of Primary Care Health Services Research, Division of Primary Health Care, University of Bristol

Professor Ian Roberts, Professor of Epidemiology & Public Health, Intervention Research Unit, London School of Hygiene and Tropical Medicine

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Professor Mark Sculpher, Professor of Health Economics, Centre for Health Economics, Institute for Research in the Social Services, University of York

Professor Martin Severs, Professor in Elderly Health Care, Portsmouth Institute of Medicine

Dr Jonathan Shapiro, Senior Fellow, Health Services Management Centre, Birmingham

Ms Kate Thomas, Deputy Director, Medical Care Research Unit, University of Sheffield

Professor Simon G Thompson, Director, MRC Biostatistics Unit, Institute of Public Health, Cambridge

Ms Sue Ziebland, Senior Research Fellow, Cancer Research UK, University of Oxford

Current and past membership details of all HTA 'committees' are available from the HTA website (www.ncchta.org)

Diagnostic Technologies & Screening Panel

Members

Chair,
Dr Ron Zimmern, Director of the Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

Ms Norma Armston, Freelance Consumer Advocate, Bolton

Professor Max Bachmann Professor Health Care Interfaces, Department of Health Policy and Practice, University of East Anglia

Professor Rudy Bilous Professor of Clinical Medicine & Consultant Physician, The Academic Centre, South Tees Hospitals NHS Trust

Dr Paul Cockcroft, Consultant Medical Microbiologist/Laboratory Director, Public Health Laboratory, St Mary's Hospital, Portsmouth

Professor Adrian K Dixon, Professor of Radiology, Addenbrooke's Hospital, Cambridge

Dr David Elliman, Consultant in Community Child Health, London

Professor Glyn Elwyn, Primary Medical Care Research Group, Swansea Clinical School, University of Wales Swansea

Dr John Fielding, Consultant Radiologist, Radiology Department, Royal Shrewsbury Hospital

Dr Karen N Foster, Clinical Lecturer, Dept of General Practice & Primary Care, University of Aberdeen

Professor Antony J Franks, Deputy Medical Director, The Leeds Teaching Hospitals NHS Trust

Mr Tam Fry, Honorary Chairman, Child Growth Foundation, London

Dr Edmund Jessop, Medical Adviser, National Specialist Commissioning Advisory Group (NSCAG), Department of Health, London

Dr Jennifer J Kurinczuk, Consultant Clinical Epidemiologist, National Perinatal Epidemiology Unit, Oxford

Dr Susanne M Ludgate, Medical Director, Medical Devices Agency, London

Dr William Rosenberg, Senior Lecturer and Consultant in Medicine, University of Southampton

Dr Susan Schonfield, CPHM Specialised Services Commissioning, Croydon Primary Care Trust

Dr Margaret Somerville, Director of Public Health, Teignbridge Primary Care Trust

Professor Lindsay Wilson Turnbull, Scientific Director, Centre for MR Investigations & YCR Professor of Radiology, University of Hull

Professor Martin J Whittle, Head of Division of Reproductive & Child Health, University of Birmingham

Dr Dennis Wright, Consultant Biochemist & Clinical Director, Pathology & The Kennedy Galton Centre, Northwick Park & St Mark's Hospitals, Harrow

Pharmaceuticals Panel

Members

Chair,
Dr John Reynolds, Clinical Director, Acute General Medicine SDU, Oxford Radcliffe Hospital

Professor Tony Avery, Professor of Primary Health Care, University of Nottingham

Professor Stirling Bryan, Professor of Health Economics, Health Services Management Centre, University of Birmingham

Mr Peter Cardy, Chief Executive, Macmillan Cancer Relief, London

Dr Christopher Cates, GP and Cochrane Editor, Bushey Health Centre

Professor Imti Choonara, Professor in Child Health, University of Nottingham, Derbyshire Children's Hospital

Mr Charles Dobson, Special Projects Adviser, Department of Health

Dr Robin Ferner, Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham

Dr Karen A Fitzgerald, Pharmaceutical Adviser, Bro Taf Health Authority, Cardiff

Mrs Sharon Hart, Managing Editor, *Drug & Therapeutics Bulletin*, London

Dr Christine Hine, Consultant in Public Health Medicine, Bristol South & West Primary Care Trust

Professor Stan Kaye, Professor of Medical Oncology, Consultant in Medical Oncology/Drug Development, The Royal Marsden Hospital

Ms Barbara Meredith, Project Manager Clinical Guidelines, Patient Involvement Unit, NICE

Dr Frances Rotblat, CPMP Delegate, Medicines Control Agency, London

Professor Jan Scott, Professor of Psychological Treatments, Institute of Psychiatry, University of London

Mrs Katrina Simister, New Products Manager, National Prescribing Centre, Liverpool

Dr Richard Tiner, Medical Director, Association of the British Pharmaceutical Industry

Dr Helen Williams, Consultant Microbiologist, Norfolk & Norwich University Hospital NHS Trust

Therapeutic Procedures Panel

Members

Chair,
Professor Bruce Campbell,
Consultant Vascular and General Surgeon, Royal Devon & Exeter Hospital

Dr Mahmood Adil, Head of Clinical Support & Health Protection, Directorate of Health and Social Care (North), Department of Health, Manchester

Dr Aileen Clarke, Reader in Health Services Research, Public Health & Policy Research Unit, Barts & the London School of Medicine & Dentistry, Institute of Community Health Sciences, Queen Mary, University of London

Mr Matthew William Cooke, Senior Clinical Lecturer and Honorary Consultant, Emergency Department, University of Warwick, Coventry & Warwickshire NHS Trust, Division of Health in the Community, Centre for Primary Health Care Studies, Coventry

Dr Carl E Counsell, Senior Lecturer in Neurology, University of Aberdeen

Dr Keith Dodd, Consultant Paediatrician, Derbyshire Children's Hospital

Professor Gene Feder, Professor of Primary Care R&D, Barts & the London, Queen Mary's School of Medicine and Dentistry, University of London

Professor Paul Gregg, Professor of Orthopaedic Surgical Science, Department of Orthopaedic Surgery, South Tees Hospital NHS Trust

Ms Bec Hanley, Freelance Consumer Advocate, Hurstpierpoint

Ms Maryann L. Hardy, Lecturer, Division of Radiography, University of Bradford

Professor Alan Horwich, Director of Clinical R&D, The Institute of Cancer Research, London

Dr Phillip Leech, Principal Medical Officer for Primary Care, Department of Health, London

Dr Simon de Lusignan, Senior Lecturer, Primary Care Informatics, Department of Community Health Sciences, St George's Hospital Medical School, London

Dr Mike McGovern, Senior Medical Officer, Heart Team, Department of Health, London

Professor James Neilson, Professor of Obstetrics and Gynaecology, Dept of Obstetrics and Gynaecology, University of Liverpool, Liverpool Women's Hospital

Dr John C Pounsford, Consultant Physician, North Bristol NHS Trust

Dr Vimal Sharma, Consultant Psychiatrist & Hon Snr Lecturer, Mental Health Resource Centre, Victoria Central Hospital, Wirrall

Dr L David Smith, Consultant Cardiologist, Royal Devon & Exeter Hospital

Professor Norman Waugh, Professor of Public Health, University of Aberdeen

Expert Advisory Network

Members

Professor Douglas Altman,
Director of CSM & Cancer Research UK Med Stat Gp, Centre for Statistics in Medicine, University of Oxford, Institute of Health Sciences, Headington, Oxford

Professor John Bond,
Director, Centre for Health Services Research, University of Newcastle upon Tyne, School of Population & Health Sciences, Newcastle upon Tyne

Mr Shaun Brogan,
Chief Executive, Ridgeway Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Office of the Chief Executive. Trust Headquarters, Altnagelvin Hospitals Health & Social Services Trust, Altnagelvin Area Hospital, Londonderry

Ms Tracy Bury,
Project Manager, World Confederation for Physical Therapy, London

Mr John A Cairns,
Professor of Health Economics, Health Economics Research Unit, University of Aberdeen

Professor Iain T Cameron,
Professor of Obstetrics and Gynaecology and Head of the School of Medicine, University of Southampton

Dr Christine Clark,
Medical Writer & Consultant Pharmacist, Rossendale

Professor Collette Mary Clifford,
Professor of Nursing & Head of Research, School of Health Sciences, University of Birmingham, Edgbaston, Birmingham

Professor Barry Cookson,
Director, Laboratory of Healthcare Associated Infection, Health Protection Agency, London

Professor Howard Stephen Cuckle,
Professor of Reproductive Epidemiology, Department of Paediatrics, Obstetrics & Gynaecology, University of Leeds

Professor Nicky Cullum,
Director of Centre for Evidence Based Nursing, University of York

Dr Katherine Darton,
Information Unit, MIND – The Mental Health Charity, London

Professor Carol Dezateux,
Professor of Paediatric Epidemiology, London

Mr John Dunning,
Consultant Cardiothoracic Surgeon, Cardiothoracic Surgical Unit, Papworth Hospital NHS Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon, Gloucestershire Royal Hospital, Gloucester

Professor Martin Eccles,
Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne

Professor Pam Enderby,
Professor of Community Rehabilitation, Institute of General Practice and Primary Care, University of Sheffield

Mr Leonard R Fenwick,
Chief Executive, Newcastle upon Tyne Hospitals NHS Trust

Professor David Field,
Professor of Neonatal Medicine, Child Health, The Leicester Royal Infirmary NHS Trust

Mrs Gillian Fletcher,
Antenatal Teacher & Tutor and President, National Childbirth Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine, Department of Medicine, University of Birmingham, Queen Elizabeth Hospital, Edgbaston, Birmingham

Ms Grace Gibbs,
Deputy Chief Executive, Director for Nursing, Midwifery & Clinical Support Servs, West Middlesex University Hospital, Isleworth

Dr Neville Goodman,
Consultant Anaesthetist, Southmead Hospital, Bristol

Professor Alastair Gray,
Professor of Health Economics, Department of Public Health, University of Oxford

Professor Robert E Hawkins,
CRC Professor and Director of Medical Oncology, Christie CRC Research Centre, Christie Hospital NHS Trust, Manchester

Professor F D Richard Hobbs,
Professor of Primary Care & General Practice, Department of Primary Care & General Practice, University of Birmingham

Professor Allen Hutchinson,
Director of Public Health & Deputy Dean of ScHARR, Department of Public Health, University of Sheffield

Dr Duncan Keeley,
General Practitioner (Dr Burch & Ptnrs), The Health Centre, Thame

Dr Donna Lamping,
Research Degrees Programme Director & Reader in Psychology, Health Services Research Unit, London School of Hygiene and Tropical Medicine, London

Mr George Levvy,
Chief Executive, Motor Neurone Disease Association, Northampton

Professor James Lindesay,
Professor of Psychiatry for the Elderly, University of Leicester, Leicester General Hospital

Professor Rajan Madhok,
Medical Director & Director of Public Health, Directorate of Clinical Strategy & Public Health, North & East Yorkshire & Northern Lincolnshire Health Authority, York

Professor David Mant,
Professor of General Practice, Department of Primary Care, University of Oxford

Professor Alexander Markham,
Director, Molecular Medicine Unit, St James's University Hospital, Leeds

Dr Chris McCall,
General Practitioner, The Hadleigh Practice, Castle Mullen

Professor Alistair McGuire,
Professor of Health Economics, London School of Economics

Dr Peter Moore,
Freelance Science Writer, Ashtead

Dr Andrew Mortimore,
Consultant in Public Health Medicine, Southampton City Primary Care Trust

Dr Sue Moss,
Associate Director, Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton

Professor Jon Nicholl,
Director of Medical Care Research Unit, School of Health and Related Research, University of Sheffield

Mrs Julietta Patnick,
National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry, University Mental Health Group, Royal South Hants Hospital, Southampton

Professor Chris Price,
Visiting Chair – Oxford, Clinical Research, Bayer Diagnostics Europe, Cirencester

Ms Marianne Rigge,
Director, College of Health, London

Dr Eamonn Sheridan,
Consultant in Clinical Genetics, Genetics Department, St James's University Hospital, Leeds

Dr Ken Stein,
Senior Clinical Lecturer in Public Health, Director, Peninsula Technology Assessment Group, University of Exeter

Professor Sarah Stewart-Brown,
Director HSRU/Honorary Consultant in PH Medicine, Department of Public Health, University of Oxford

Professor Ala Szczepura,
Professor of Health Service Research, Centre for Health Services Studies, University of Warwick

Dr Ross Taylor,
Senior Lecturer, Department of General Practice and Primary Care, University of Aberdeen

Mrs Joan Webster,
Consumer member, HTA – Expert Advisory Network

Feedback

The HTA Programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278