Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial hyperplasia

Authors' objectives
The authors investigated the accuracy of out-patient endometrial biopsy in diagnosing endometrial hyperplasia in women with abnormal uterine bleeding.

Searching
MEDLINE and EMBASE were searched from 1980 to 1999 using MeSH terms with 'diagnosis', in combination with textwords 'endometrial biopsy' and 'diagnosis'. The reference lists of identified reviews and primary studies were also checked. Studies published in any language were eligible for inclusion. The manufacturers of out-patient endometrial biopsy devices were contacted.

Study selection
Study designs of evaluations included in the review
Prospective observational studies and comparative cross-sectional studies were eligible for inclusion.

Specific interventions included in the review
Studies of out-patient endometrial biopsy devices were eligible for inclusion. The included studies assessed the Pipelle device (3 studies), Gynoscan (1 study), Accurette and Vabra Aspiration devices (1 study), and Novak Curette and Vabra Aspiration (1 study).

Reference standard test against which the new test was compared
Studies using endometrial histology obtained by in-patient sampling as the reference standard were eligible for inclusion.

Participants included in the review
Studies of women with abnormal pre- or postmenopausal uterine bleeding were eligible for inclusion. Postmenopausal women made up 25% of the review population.

Outcomes assessed in the review
The primary outcome measure was the accuracy of diagnosing endometrial hyperplasia. The secondary measures were device failures and the rates of inadequate specimens for histological assessment.

How were decisions on the relevance of primary studies made?
Two reviewers independently screened titles and full papers for inclusion. Any disagreements were resolved by consensus, otherwise a third reviewer was consulted.

Assessment of study quality
The authors constructed their own quality checklist based on existing quality criteria. The list included the following components: study design; method of population sampling; biopsy technique description; blinding of biopsy results; and description of the outcomes and population. Two reviewers applied the validity criteria independently and resolved any disagreements by consensus. The extent of agreement between the reviewers was calculated.

Data extraction
The authors did not state how many reviewers performed the data extraction. The results were entered into 2x2 tables of out-patient biopsy results (positive or negative for hyperplasia) and the results of the reference standard histology (hyperplasia or non-hyperplasia). When hyperplasia was present, the data were stratified by the presence or absence of atypical cells and further 2x2 tables were constructed.
Data were also collected for endometrial cancer or pre-cancer (complex or atypical hyperplasia). The out-patient biopsy result (positive or negative for endometrial cancer plus pre-cancer) and the results for the reference standard histology (positive or negative for endometrial cancer plus pre-cancer) were entered into 2x2 tables.

The true-positive rate, false-positive rate and likelihood ratios (LRs) were calculated for each study and for each subgroup analysis.

Unsuccessful sampling was termed as either failed procedures or histologically inadequate specimens. Failure rates were recorded but were not entered into 2x2 tables. Inadequate specimens were included in a sensitivity analysis along with negative results.

**Methods of synthesis**

How were the studies combined?
The LRs were pooled, weighted by the inverse variance.

A funnel plot of the diagnostic odds ratios against the corresponding precision of effect was used to investigate publication bias.

How were differences between studies investigated?
The Breslow-Day test was used to assess heterogeneity between the study results. Subgroup analyses were conducted to investigate potential sources of heterogeneity. Population, index test, outcome and study quality were used to stratify the studies.

**Results of the review**

Six studies of eight diagnostic evaluations (881 women) were included. Four were classified as prospective studies, while the other two were retrospective studies.

The observer agreement for classifying quality criteria ranged from 96 to 100%; perfect agreement (kappa 1.0) was obtained on 5 items. All of the included studies were given a level 4 quality relating; this corresponded to an independent, non-blind comparison with reference standard among an appropriate population of non-consecutive patients or confined to a narrow population of study patients.

The pooled LR for positive endometrial hyperplasia for all devices was 12.0 (95% confidence interval, CI: 7.8, 18.6) and the negative LR was 0.2 (95% CI: 0.1, 0.3).

The pooled LRs of the 3 studies that used the Pipelle device were 9.9 (95% CI: 5.5, 17.6) for positive results and 0.5 (95% CI: 0.4, 0.6) for negative results.

The pooled LRs of the 2 studies that used the Vabra Aspirator device were 12.6 (95% CI: 5.6, 28.1) for positive results and 0.1 (95% CI: 0.01, 0.5) for negative results.

The pooled LR was 10.7 (95% CI: 6.1, 18.9) for positive endometrial hyperplasia without atypia and 0.45 (95% CI: 0.35, 0.57) for negative endometrial hyperplasia without atypia. The corresponding LRs for endometrial hyperplasia with atypia were 34 (95% CI: 4.2, 277.4) and 0.3 (95% CI: 0.07, 1.7), respectively.

The pooled LR for positive endometrial cancer with or without complex or atypical endometrial hyperplasia was 95.1 (95% CI: 41.2, 219.4); the pooled LR for negative endometrial cancer with or without complex or atypical endometrial hyperplasia was 0.1 (95% CI: 0.07, 0.3).

There was significant heterogeneity across all studies, as confirmed by the Breslow-Day test (P=0.003). When stratified by study quality, the sensitivity analyses showed an effect on the heterogeneity of diagnostic performance. Prospective recruitment (P=0.3), clear population details (P=0.3), and the exclusive use of one particular reference standard (P=0.3) removed significant heterogeneity of diagnostic test performance.
The failure rate for out-patient biopsy was 33 out of 881, which was 4% of all attempted biopsies (95% CI: 3, 5). The failure rate was lower in a sample restricted to postmenopausal women (25 out of 149, i.e. 2%, 95% CI: 1, 2.4) than in all women combined.

The Spearman rank correlation found there to be a possibility of publication bias (r=0.83, P=0.02)

**Authors' conclusions**
Out-patient endometrial biopsy has modest accuracy in diagnosing endometrial hyperplasia. Therefore additional endometrial assessment should be undertaken, especially if symptoms persist or intra-uterine structural abnormalities are suspected. The authors also concluded that the more clinically significant the endometrial pathology, the greater the accuracy of the test.

**CRD commentary**
The review question was clear and used explicit inclusion criteria. A search for published studies in all languages was performed, although no attempt was made to include unpublished literature. The authors assessed the potential for publication bias and found it to be likely. Two reviewers independently screened the titles, abstracts and full papers for inclusion and assessed the validity of the studies; this should reduce the chance of selection bias. A thorough quality assessment was performed and used in the assessment of differences between the studies. The pooling of all LRs in the presence of heterogeneity might not have been appropriate, especially as the authors did not mention whether they used a random-effects or fixed-effect model. The authors' conclusions are vague as the sensitivity of out-patient endometrial biopsy was generally poor, except when identifying cancer.

**Implications of the review for practice and research**
The authors did not state any implications for practice or further research.

**Funding**
University of Birmingham Interdisciplinary Research Fund; Birmingham Women's Hospital Research and Development Programme.

**Bibliographic details**

**PubMedID**
11531627

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Ambulatory Care; Biopsy /standards; Endometrial Hyperplasia /pathology; Endometrium /pathology; Female; Humans

**AccessionNumber**
12001002356

**Date bibliographic record published**
31/03/2005

**Date abstract record published**
31/03/2005
Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.