Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities

Smith-Bindman R, Kerlikowske K, Feldstein V A, Subak L, Scheidler J, Segal M, Brand R, Grady D

Authors' objectives
To determine the accuracy of endovaginal ultrasound (EVUS) in detecting endometrial disease in postmenopausal women with vaginal bleeding according to hormone replacement use.

Searching
MEDLINE was searched from January 1966 to November 1996 using the following key terms: 'endometrium', 'endometrial cancer', 'ultrasound', 'endovaginal ultrasound', 'transvaginal', 'transvaginal ultrasound', 'vaginal bleeding', 'dilation and curettage' and 'endometrial biopsy'. The bibliographies of identified articles were searched manually. No language restrictions were applied.

Study selection
Study designs of evaluations included in the review
Studies in which EVUS was prospectively evaluated were included. The exclusion criteria included retrospective studies, studies with selective histological sampling, and studies not reporting endovaginal thickness measurements. Review articles, letters, case reports and comments were also excluded. Published and unpublished studies were included.

Specific interventions included in the review
Studies that evaluated EVUS measurement of endometrial thickness prior to obtaining endometrial tissue were eligible for inclusion. Studies measuring endometrial thickness following tissue sampling were excluded.

Reference standard test against which the new test was compared
Studies using histological evaluation of endometrial tissue (endometrial biopsy, dilation and curettage, or hysterectomy) as the reference standard to detect endometrial disease were eligible for inclusion.

Participants included in the review
The included studies were of women presenting with postmenopausal bleeding. The mean age of the participants was 61 years and 94% were symptomatic with vaginal bleeding. The following groups were excluded: asymptomatic women; women receiving tamoxifen; and women with known endometrial cancer, cervical cancer or cervical polyps. Studies reporting pooled data for pre- and postmenopausal women were also excluded.

Outcomes assessed in the review
Studies where crude data could not be extracted or obtained through contacting the authors were excluded. The outcome measures calculated in the review were the sensitivity and specificity.

How were decisions on the relevance of primary studies made?
Two reviewers independently selected and reviewed the articles.

Assessment of study quality
The authors did not state that they assessed validity.

Data extraction
For each outcome, two authors extracted the following data: the numbers of true-positive, false-positive, true-negative and false-negative cases using all reported thickness thresholds (3, 4, 5, 6, 7, 8, 9 and 10 mm); the mean endometrial thickness among women with atrophic endometrium, hyperplasia or polyps or cancer; the mean age of the participants; the number using hormone replacement therapy (HRT); the number with vaginal bleeding; frequency of the ultrasound...
Methods of synthesis
How were the studies combined?
For each study, the sensitivity and specificity and exact 95% confidence intervals (CIs) were calculated for all EVUS thickness measurements. Mean weighted pooled estimates of sensitivity and specificity were estimated for each threshold, for any endometrial disease and for cancer alone. Weighting was undertaken using sample size. The 95% CIs were calculated using exact methods. Positive and negative likelihood ratios were calculated for each endometrial thickness. The post-test risk of endometrial disease was calculated using the 5-mm thickness threshold and varying the pre-test risk of disease from 1 to 50%. A summary receiver operating characteristic (ROC) curve was generated using data from all thresholds using the method of Moses et al. (see Other Publications of Related Interest no.1). Separate summary ROC curves were generated for women who used HRT and those who did not. Studies in which the use of HRT was not explicitly stated were included with studies that included women using HRT.

How were differences between studies investigated?
Homogeneity was assessed by determining whether the CIs from individual studies overlapped the weighted summary point estimate. Heterogeneity was investigated by stratifying the summary estimate according to the following factors: English or non English language of publication; use of HRT; patient symptoms; whether fluid within the endometrial cavity was included or excluded in determining endometrial thickness; and speciality of the examiner. Differences between the summary ROC curves were tested for women using HRT and those who did not, and between women who used HRT and those in whom HRT was not stated. A sensitivity analysis was undertaken by including studies excluded on methodological criteria. The threshold effect was assessed, prior to pooling, by examining the correlation between sensitivity and specificity with any individual thickness measurement used to define abnormal across different studies.

Results of the review
Thirty-five studies (5,892 women) were included.

Prevalence of disease was 13% for endometrial cancer and 40% for endometrial polyps or hyperplasia.

The number of women unable to tolerate EVUS was reported in 16 studies; 14 studies reported the number of women who had a non-diagnostic EVUS.

Mean endometrial thickness according to pathology: normal, 4 mm (standard deviation, SD=1); endometrial polyp, 10 mm (SD=3); hyperplasia, 14 mm (SD=4); and cancer, 20 mm (SD=6).

There was no correlation between sensitivity and specificity within any of the individual thickness measurements used to define abnormal in the individual studies.

The sensitivity and specificity were reported for endometrial disease using different endovaginal thickness measurements to define an abnormal result. Using a 5-mm thickness threshold: the sensitivity for endometrial disease (1,306 women) was 92% (95% CI: 90, 93) and the specificity (2,137 women) was 81% (95% CI: 79, 83); the sensitivity for cancer (457 women) was 96% (95% CI: 94, 98) and the specificity (2,986 women) was 61% (95% CI: 59, 63).

The accuracy of detecting endometrial disease, stratified by HRT status, was reported for different endovaginal thickness thresholds. Using a 5-mm threshold, with no HRT, the sensitivity (423 women) was 95% (95% CI: 93, 97), the specificity (593 women) 92% (95% CI: 90, 94), the positive likelihood 11.9, and the negative likelihood 0.05. With HRT, the sensitivity (883 women) was 91% (95% CI: 89, 93), the specificity (1,544 women) 77% (95% CI: 75, 79), the positive likelihood 4.0, and the negative likelihood 0.12.

Homogeneity. For sensitivity, using a 5-mm thickness threshold, the studies were homogeneous for cancer, while 2 out of 20 were heterogeneous for any endometrial disease. For specificity, the results were inconsistent across the studies.
Using a 5-mm thickness threshold, 7 out of 20 studies were heterogeneous for cancer and 8 out of 20 were heterogeneous for any endometrial disease. The estimates for specificity were less heterogeneous after stratifying by the use of HRT, with heterogeneity improving among women not using HRT but remaining in women using HRT. Stratification by other factors did not improve consistency across the studies.

The results confirmed the trade-off between sensitivity and specificity obtained by changing the threshold thickness used to define abnormality. At 3 mm, the sensitivity was 98% and the specificity 38%; at 10 mm, the sensitivity was 66% and the specificity 79%. Summary ROC curves were significantly different between women using HRT and those not using HRT (P=0.02).

In the sensitivity analysis, the inclusion of excluded studies had little effect. After the exclusion of one study that contributed 25% of the patients to the pooled results, the sensitivity for endometrial disease at a 5-mm thickness threshold decreased from 92 to 88%; sensitivity for cancer alone decreased from 96 to 95%.

Cost information
The authors reported that the cost of EVUS compared favourably with endometrial biopsy in the evaluation of postmenopausal bleeding (see Other Publications of Related Interest no.2).

Authors' conclusions
Endovaginal ultrasound has a high sensitivity for detecting endometrial cancer and other endometrial disease, and can reliably identify postmenopausal women with vaginal bleeding who are highly unlikely to have significant endometrial disease, so that endometrial sampling may be unnecessary.

CRD commentary
The aims and the inclusion criteria were stated clearly. The methods used to select the studies and extract the data were described, and the results were clearly presented. The methods used to pool the data from the primary studies were robust, and heterogeneity was assessed and investigated. The discussion considered potential causes of heterogeneity that were not investigated in the review and the possibility of publication bias.

By limiting the literature search to one database some relevant studies may have been omitted. Validity was not formally assessed, leaving open the possibility of bias arising from methodological flaws in the included studies. The authors undertook a sensitivity analysis relating to studies excluded on the basis of methodological criteria. However, the nature of these studies was unclear.

The conclusions appear to be supported by the evidence presented.

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

Bibliographic details

PubMedID
9809732

Original Paper URL
http://jama.ama-assn.org/

Other publications of related interest

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Biopsy; Endometrial Neoplasms /diagnosis /ultrasound; Endometrium /pathology /ultrasound; Endosonography; Female; Hormone Replacement Therapy; Humans; Likelihood Functions; Postmenopause; ROC Curve; Sensitivity and Specificity; Uterine Hemorrhage

**AccessionNumber**
11998009095

**Date bibliographic record published**
30/09/2004

**Date abstract record published**
30/09/2004

**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.