A systematic review of transvaginal ultrasonography, sonohysterography and hysteroscopy for the investigation of abnormal uterine bleeding in premenopausal women

Farquhar C, Ekeroma A, Furness S, Arroll B

CRD summary
This systematic review evaluated three methods for investigating abnormal uterine bleeding in women who had not yet reached the menopause. All three methods were found to be moderately accurate for detecting cancer, hyperplasia and fibroids, with transvaginal ultrasound performing the worst of the three for fibroids. The review had some methodological limitations and little was reported on discomfort experienced during the procedures.

Authors' objectives
To review the effectiveness of transvaginal ultrasonography (TVUS), sonohysterography and diagnostic hysteroscopy with biopsy for the investigation of abnormal uterine bleeding in premenopausal women.

Searching
MEDLINE and EMBASE were searched from 1980 to July 2001; the search terms were reported. The reference lists of known reviews and primary studies were checked. Only English language studies were included. Manufacturers were not contacted.

Study selection
Study designs of evaluations included in the review
Comparative studies of diagnostic tests and a reference test were eligible for inclusion. All included studies were prospective; no further details of the study designs were given.

Specific interventions included in the review
Studies of TVUS, transabdominal or transvaginal sonohysterography, or diagnostic hysteroscopy with biopsy, for the investigation of abnormal uterine bleeding in premenopausal women, were eligible for inclusion.

Reference standard test against which the new test was compared
Studies that used histopathology in combination with hysteroscopy or hysterectomy as the reference standard were eligible for inclusion. Studies of hysteroscopy as the intervention were excluded if they were not compared with a reference standard such as operative hysteroscopy or histopathology at hysterectomy. Studies of different endometrial sampling techniques were excluded.

Participants included in the review
Studies of premenopausal women with menstrual symptoms suggestive of underlying uterine pathology were eligible for inclusion. The participants in the majority of the included studies had abnormal uterine bleeding. Only studies with at least 60% premenopausal women and at least 50% women with menstrual disorders were eligible for inclusion.

Outcomes assessed in the review
The outcomes of interest were endometrial hyperplasia, endometrial polyps, submucous fibroids, technical failures, patient discomfort and adverse events. Studies were only included if it was possible to calculate sensitivity and specificity and form 2x2 tables from the results presented.

How were decisions on the relevance of primary studies made?
One author assessed the search results and another checked them. Full text articles of all provisionally included articles were retrieved and assessed against the inclusion criteria. Any disagreements were resolved by consensus between all authors.
Assessment of study quality
The authors used a checklist derived from published literature to assess validity. The criteria used were: blind independent assessment of the new and reference tests by two assessors; comparison of TVUS or sonohysterography with either a reference test of operative hysteroscopy or histopathology at hysterectomy; an appropriate spectrum of patients, defined as women with abnormal bleeding but not only those scheduled for surgery, with no postmenopausal women. Studies were graded A to D: A meant all three criteria were met, B two criteria were met, C one criteria was met and D no criteria were met. The authors did not state how the papers were assessed for validity, or many reviewers performed the validity assessment.

Data extraction
Outcome data from each study were extracted as 2x2 tables. Sensitivity, specificity, and positive and negative likelihood ratios (LRs) with 95% confidence intervals (CIs) were calculated for the outcomes endometrial hyperplasia and carcinoma, submucous fibroids, and any intrauterine pathology identified. Data from premenopausal women were used where these were presented separately; if they could not be separated, combined data from pre- and postmenopausal women were used.

Methods of synthesis
How were the studies combined?
Sensitivities and specificities were calculated. Pooled positive and negative LRs were calculated using a fixed-effect model. Where the data in a 2x2 table of a study included 0 then 1 was added to that cell. Failure rates were recorded but were excluded from the 2x2 tables.

How were differences between studies investigated?
Statistical heterogeneity of the studies was assessed using the chi-squared test. The chi-squared test was taken to indicate significant heterogeneity where the P-value was less than 0.05. A subgroup analysis was performed according to variation in study quality. Pooled LRs were not presented where significant heterogeneity was found.

Results of the review
Nineteen studies involving 2,917 women were included: 15 studies of premenopausal women (n=2,509) and 4 studies with combined data from pre- and postmenopausal women (n=408)).

The results presented in the paper were not always consistent between the text and tables. The results presented here were taken from the text.

Five studies were graded A, nine were graded B and five were graded C.

TVUS versus histopathology combined with hysteroscopy or hysterectomy.

Detection of any intrauterine pathology (10 studies, 4 grade A).

The results were not pooled because of heterogeneity. The sensitivity ranged from 46 to 100%, while the specificity ranged from 12 to 100%. The positive LRs ranged from 1.05 to 51.56 and the negative LRs from 0.07 to 0.79.

Detection of submucous fibroids (9 studies, 1 grade A). The results were not pooled because of heterogeneity. The sensitivity ranged from 21 to 100%, while the specificity ranged from 53 to 100%. The positive LRs ranged from 1.61 to 62.25 and the negative LR from 0.03 to 0.80.

Detection of endometrial hyperplasia or carcinoma (7 studies, 3 grade A).

The results were not pooled because of heterogeneity. The sensitivity ranged from 33 to 100%, while the specificity ranged from 79 to 99%. The positive LRs ranged from 2.59 to 679 and the negative LRs from 0.04 to 1.00. Cut-offs of 8 to 14 mm were used in the included studies.
When only studies graded A for quality were included, heterogeneity was still present for all outcomes except negative LRs for the outcome of endometrial hyperplasia or carcinoma (LR 0.05, 95% CI: 0.03, 0.10).

Sonohysterography versus histopathology.

Detection of all intrauterine pathologies (11 studies).

Heterogeneity was present for the following (data not pooled): sensitivity (range: 85 to 100%), specificity (range: 81 to 100%) and positive LRs (range: 1.96 to 80.0). The pooled negative LR was 0.12 (95% CI: 0.08, 0.18).

Detection of submucous fibroids (7 studies).

Heterogeneity was present for the following (data not pooled): sensitivity (range: 57 to 100%), specificity (range: 96 to 100%) and negative LRs (range: 0.06 to 0.47). The pooled positive LR was 29.6 (95% CI: 17.77, 49.60).

Detection of endometrial hyperplasia or carcinoma (4 studies).

The results were not pooled because of heterogeneity. The sensitivity ranged from 29 to 80%, while the specificity ranged from 82 to 100%. The positive LRs ranged from 1.55 to 70.40 and the negative LRs from 0.14 to 0.88.

A sensitivity analysis was not possible as only one study was graded A for quality.

Diagnostic hysteroscopy and biopsy versus reference standard.

Detection of all intrauterine pathologies (3 studies). Heterogeneity was present for the following (data not pooled): sensitivity (range: 90 to 97%), specificity (range: 62 to 93%) and positive LRs (range: 2.55 to 14.56). The pooled negative LR was 0.07 (95% CI: 0.04, 0.15).

Detection of submucous fibroids (4 studies).

Heterogeneity was present for the following (data not pooled): sensitivity (range: 53 to 100%), specificity (range: 97 to 100%) and negative LRs (range: 0.08 to 0.48). The pooled positive LR was 29.43 (95% CI: 13.26, 65.33).

Detection of endometrial hyperplasia or carcinoma (3 studies).

Heterogeneity was present for the following (data not pooled): sensitivity (range: 90 to 100%) and specificity (range: 97 to 100%). The pooled positive LR was 92.84 (95% CI: 47.0, 111.7) and the pooled negative LR was 0.05 (95% CI: 0.02, 0.12).

Patient discomfort.

One study of TVUS reported no cases of discomfort, while a second reported that 2% of patients found the procedure unpleasant and 40% felt slight pelvic pain.

Four studies using sonohysterography reported that some patients experienced discomfort. A fifth study reported that 13% of patients found the procedure unpleasant and 53% felt slight pelvic pain. In a sixth study, 97 of the 130 women reported no pain or acceptable discomfort, and one had barely tolerable pain. In a seventh study, 4 of the 50 patients had severe pain and all patients reported some pain.

One study using hysteroscopy reported that 13 of the 793 cases were not completed because of intolerance of the procedure, while 3.6% of women reported unacceptable pain. Another study found 50 of the 130 women had either tolerable or barely tolerable discomfort with this procedure.

Information on morbidity was reported infrequently.

Authors' conclusions
All three diagnostic tests were moderately accurate in detecting intrauterine pathology, but there was considerable variation between the studies. Sonohysterography and hysteroscopy performed better than TVUS in detecting submucous fibroids.

CRD commentary
The inclusion criteria for the index and reference tests, participants and outcomes were clearly stated, although it was unclear which study designs were included. More than one electronic database was searched for primary studies, but no attempt to locate unpublished studies was made and inclusion was restricted to English language publications; this might have resulted in some studies being missed. Details of the review process (how many reviewers assessed validity and extracted the data) were not reported, so it was not possible to judge whether bias might have occurred at this stage, although more than one reviewer independently screened studies for inclusion, which minimises bias. The validity of the included studies was assessed using a checklist derived from the published literature and used in a sensitivity analysis. A meta-analysis was performed and studies were only pooled where no significant heterogeneity existed, which seems appropriate. The authors' conclusions did not mention patient discomfort, which seems to be an important element accompanying the more invasive tests.

Implications of the review for practice and research
Practice: The authors did not report any implications for practice.
Research: The authors stated that randomised trials need to be undertaken with specific outcomes in mind. In particular, whether test results aided in planning patient management; whether the test led to the avoidance of other medical procedures; whether test information changed therapeutic choices; and clinical outcomes. The reporting of inter-rater variability should be encouraged. Future research should examine the clinical usefulness of the tests in a clinical setting.

Bibliographic details

PubMedID
12780419

Indexing Status
Subject indexing assigned by NLM

MeSH
Diagnosis, Differential; False Negative Reactions; False Positive Reactions; Female; Humans; Hysteroscopy; Premenopause; Sensitivity and Specificity; Ultrasonography /methods; Uterine Diseases /complications /diagnosis /ultrasonography; Uterine Hemorrhage /ultrasonography; Vagina /ultrasonography

AccessionNumber
12003001324

Date bibliographic record published
31/10/2005

Date abstract record published
31/10/2005

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.