Diagnostic hysteroscopy in abnormal uterine bleeding: a systematic review and meta-analysis
van Dongen H, de Kroon C D, Jacobi C E, Trimbos J B, Jansen F W

CRD summary
This review found that diagnostic hysteroscopy is accurate and feasible for the diagnosis of intra-uterine abnormalities. Although generally well conducted and clearly reported, limitations in the literature search and analysis mean that these conclusions should be interpreted with extreme caution.

Authors’ objectives
To determine the accuracy of diagnostic hysteroscopy for the evaluation of intra-uterine abnormalities in women with abnormal uterine bleeding.

Searching
MEDLINE, EMBASE, Current Contents, Science Citation Index and the Cochrane Library were searched from inception to January 2006 with no language restrictions. The search terms were reported; these included a diagnostic filter. The reference lists of retrieved articles were screened for additional relevant studies.

Study selection
Study designs of evaluations included in the review
Inclusion criteria were not specified in terms of the study design.

Specific interventions included in the review
Studies of diagnostic hysteroscopy were eligible for inclusion. Studies of hysteroscopy specifically designed for therapy were excluded.

Reference standard test against which the new test was compared
Studies in which the reference standard was histology, collected using guided biopsy during hysteroscopy, operative hysteroscopy or hysterectomy, were eligible for inclusion. The reference standards reported in the included studies were guided biopsy, dilatation and curettage, hysterectomy and operative hysteroscopy. Some studies used several different reference standards.

Participants included in the review
Studies of pre- and postmenopausal women of whom at least 70% had abnormal uterine bleeding were eligible for inclusion. Studies were excluded if more than 5% of women used tamoxifen or if the primary reason for hysteroscopy was for fertility problems.

Outcomes assessed in the review
Studies had to report data on diagnostic accuracy.

How were decisions on the relevance of primary studies made?
Two reviewers screened studies for inclusion.

Assessment of study quality
Studies were assessed for methodological quality based on the following criteria, where the numbers in brackets indicate the maximum points awarded for each item: reference standard of hysterectomy (2); one reference standard within the study (1); definition of polyp and fibroid (1); interpretation of reference blinded to index test (1); avoidance of verification bias (1); prospective study design (1); in-patient setting (1); failure rate reported (1); inclusion criteria reported (1); exclusion criteria reported (1); information on procedure available (1); age (1) and menopausal state (1) of study population reported. Studies scoring at least 10 were considered to be of a high quality, studies scoring 6 to 9 were of average quality, and studies scoring 5 or less were of low quality. Two reviewers assessed study quality and any differences were resolved through consensus.
**Data extraction**
Two reviewers independently extracted data as 2x2 tables of test performance. Where necessary, authors were contacted for further information. Separate 2x2 tables were constructed for benign intra-uterine disorders such as endometrial polyps and submucous myomas. The sensitivity and specificity, together with their 95% confidence intervals (CIs), were calculated for each study. Hysteroscopic procedures in which a final diagnosis was not made were categorised as failed procedures and recorded. Complication rates were also recorded.

**Methods of synthesis**
How were the studies combined?
Fixed-effect models were used to pool sensitivity, specificity and likelihood ratios. If significant heterogeneity was present (p>0.05) then a random-effects model was used. Pooled likelihood ratios were used to transform the pre-test probability of intra-uterine abnormalities in the included studies to give post-test probabilities (predictive values). The correlation between sensitivity and specificity was investigated using Spearman's correlation coefficient; if this was less than -0.25 then a summary receiver operating characteristic curve was constructed.

How were differences between studies investigated?
Heterogeneity was assessed using the chi-squared and I-squared statistics. A subgroup analysis based on quality, menopausal state, time, setting and performance was carried out.

**Results of the review**
Seventeen studies (4,208 procedures) were included.

The quality scores ranged from 4 to 13 out of a maximum score of 14. The proportion of patients who received verification was 100% in 14 studies and 44 to 98% in the remaining 3 studies.

Diagnosis of intracavitary abnormalities (16 studies).
The sensitivity ranged from 84 to 100% and the pooled sensitivity was 94% (95% CI: 92, 95). There was no evidence of heterogeneity (p=0.37; I-squared 7.6%). The specificity ranged from 39 to 100% and the pooled specificity was 89% (95% CI: 87, 90). There was strong evidence of heterogeneity (p<0.001; I-squared 81.5%). The prevalence of uterine abnormalities was 46.6%. The only subgroup analysis in which heterogeneity was eliminated was for the 5 studies of postmenopausal bleeding. Pooled sensitivity and specificity for this subgroup were similar to the overall figures.

Endometrial polyps (12 studies).
The pooled sensitivity and specificity were 94% (95% CI: 92, 96) and 92% (95% CI: 91, 94), respectively. Ranges were not reported, neither were data on heterogeneity.

Submucous myomas (11 studies).
The pooled sensitivity and specificity were 87% (95% CI: 81, 92) and 95% (95% CI: 93, 97), respectively. Ranges were not reported, neither were data on heterogeneity.

The proportion of failed procedures was reported in 12 out of 17 studies and ranged from 0 to 17%; 7 studies reported 0 failed procedures. Complication rates were reported for 1,399 procedures and showed 16 complications: 13 vasovagal collapses, two false tracts and one perforation of the uterine wall.

**Authors' conclusions**
Diagnostic hysteroscopy is accurate and feasible for the diagnosis of intra-uterine abnormalities.

**CRD commentary**
The review addressed a focused question that was supported by clearly defined inclusion criteria. The literature search involved an appropriate range of databases, but the use of a diagnostic filter means that relevant studies are likely to have been missed. In addition, no attempts were made to locate unpublished studies and so the review may be subject to...
publication bias. Full details of the review process were reported and these included appropriate steps to minimise bias and errors. The synthesis of the results suffered from a number of limitations: the methods focused on identifying subgroups in which there was no heterogeneity rather than attempting to identify explanations for the observed heterogeneity. Considerable heterogeneity remained in estimates of specificity for the overall analysis, thus these pooled data should be interpreted with extreme caution. Data on heterogeneity and ranges in sensitivity and specificity were not reported for the subgroup analyses of endometrial polyps and submucous myomas, making the pooled data difficult to interpret. Limitations in the search and analysis mean that the authors’ conclusions should be interpreted with extreme caution.

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

**Funding**
Not stated.

**Bibliographic details**

**PubMedID**
17516956

**DOI**
10.1111/j.1471-0528.2007.01326.x

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Endometrium /pathology; Feasibility Studies; Female; Humans; Hysteroscopy /methods /standards; Polyps /diagnosis; Sensitivity and Specificity; Uterine Hemorrhage /diagnosis

**AccessionNumber**
12007001970

**Date bibliographic record published**
01/04/2008

**Date abstract record published**
01/09/2008

**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.