The diagnostic work-up of women with postmenopausal bleeding
van Hanegem, N.

Citation for published version (APA):
van Hanegem, N. (2015). The diagnostic work-up of women with postmenopausal bleeding
Chapter 2

Diagnostic evaluation of the endometrium in postmenopausal bleeding: an evidence-based approach

N. van Hanegem, M.C. Breijer, K.S. Khan, T.J. Clark, M.P.M. Burger, B.W.J. Mol & A. Timmermans

* Maturitas 2011 Feb; 68(2): 155-64
Chapter 2

Abstract

Postmenopausal bleeding (PMB) is a common complaint in general gynaecological practice. Women with PMB have around a 10% chance of having endometrial cancer and therefore PMB always needs further evaluation. This article summarises the reviews on the subject and provides an overview of the use of diagnostic tools in patients with PMB. Four types of diagnostic test are described: sonographic measurement of endometrial thickness, endometrial sampling, hysteroscopy and saline infusion sonography. All four have been independently shown to be accurate in excluding endometrial cancer. However, neither in systematic reviews nor in international guidelines is consensus found regarding the sequence in which these methods should be employed in women with PMB. For measurement of endometrial thickness in symptomatic women, a cut-off value of three millimetres is recommended, but the cost-effectiveness of this strategy has yet to be shown. Research should now focus on the incorporation of individual patient characteristics and pre-test probabilities for cancer in algorithms for the investigation of PMB, and the most cost-effective sequenced combination of the four types of tests.
Introduction

Postmenopausal bleeding (PMB), defined as blood loss occurring at least 12 months after menopause, is a common complaint in general gynaecological practice. The prevalence of PMB is approximately 10% immediately after menopause. Postmenopausal bleeding (PMB) signals endometrial cancer in around 10% of cases, or less serious conditions, such as benign endometrial polyps, in a further 20 to 40%. Endometrial cancer is the most common gynaecologic cancer and 95% of women with endometrial cancer present with PMB. Unlike ovarian cancer, endometrial cancer often presents at an early stage, when there is a possibility of curative treatment by hysterectomy (and bilateral salpingo-oophorectomy); therefore early, accurate and timely diagnosis is important. Any PMB needs further investigation.

In the past, the principal method of investigation was dilatation and curettage (D&C). To reduce the invasiveness of investigatory procedures, ultrasonography was introduced. Endometrial biopsy and hysteroscopy have now almost completely replaced D&C. The use of outpatient endometrial biopsy reduces costs in the diagnostic work-up, without affecting life expectancy. Despite many studies on the investigation of PMB, there is still no consensus on the most accurate and efficient diagnostic pathway. This article describes a systematic literature search for guidelines and systematic reviews on this subject. The aim is to recommend an evidence-based diagnostic pathway for patients with PMB.

Methods

Identification of studies

We performed a computerised MEDLINE and EMBASE search to identify all studies on the evaluation of PMB published between January 1965 and January 2010. The search was limited to human studies; language restrictions were not applied. We included systematic review articles of observational studies on the evaluation of the endometrium in women with PMB. In addition we searched for national and international guidelines on this subject. References cited in the selected reviews were checked for further relevant articles not identified by the electronic searches. We used all known synonyms for the following keywords: postmenopausal bleeding, endometrial thickness, ultrasound, hysteroscopy and biopsy. The search strategies for the two databases are detailed in Appendices A and B.
Selection criteria
This review focused on systematic reviews in which the results of the diagnostic test of interest were compared with the results of a reference standard. The following criteria were used to select articles:

1. Population of interest was women with postmenopausal uterine bleeding.
2. The four types of diagnostic test of interest were: (a) measurement of endometrial thickness by transvaginal sonography (TVS); (b) outpatient endometrial sampling; (c) saline infusion sonography (SIS); or (d) hysteroscopy (i.e. endoscopic visual interpretation).
3. The reference standard was the endometrial histological findings from inpatient endometrial sampling, D&C or hysterectomy.
4. The primary outcome measure was the accuracy with which endometrial cancer and/or hyperplasia were diagnosed.

Studies in which more than 10% of the women used Tamoxifen were excluded, because of different pathophysiology and different characteristics of the uterine cavity. For studies that included pre- and postmenopausal women we used only those calculations and conclusions concerning the latter. The systematic reviews were selected by two reviewers working independently (NvH and MCB), through assessment of the titles and abstracts of all retrieved studies. In case of disagreement the article was included for full reading and/or assessed by a third reviewer (AT).

Additionally, we identified national and international guidelines on diagnostic procedures for PMB. From every guideline we extracted the diagnostic pathway recommended and points of policy.

Quality assessment
The methodological quality of each selected paper was assessed using the Cochrane checklist for systematic reviews of diagnostic studies. This list of criteria was designed to assess the usability of a review for guideline development. Because there is no validated checklist for the quality assessment of systematic reviews, this checklist was used as an aid in the reviewers’ evaluation of the quality of the original review, but no decisions regarding the inclusion or exclusion of articles were based on this assessment.
Data extraction, analysis and interpretation

From each systematic review or meta-analysis we extracted (if available) figures for the sensitivity, specificity, likelihood ratios (LRs), the pre-test probability of endometrial cancer and/or hyperplasia and the post-test probability. If an article described both post- and premenopausal women we extracted the data relating to the postmenopausal women. We used only data from studies evaluated as high quality by the reviewers.

The LR indicates by how much a specific test raises or lowers the probability of having endometrial pathology. An LR of 1 indicates that the test has no predictive value for the outcome of interest. The higher an LR is above 1, the larger is the probability of pathology. An LR of less than 1 indicates that a negative test result is more likely to be true. In the present study, for a rating of ‘high’ diagnostic accuracy, the LR had to be over 10 for a positive test result or less than 0.1 for a negative result. In this article, we use the LR of a negative test result (LR−) for reviews on the use of TVS, because we are interested in its accuracy in excluding endometrial cancer. For endometrial sampling and hysteroscopy we are interested in diagnosing cancer, so we used the LR of a positive test (LR+) and calculated the post-test probability of a positive test. When an article did not report post-test probability, we used Bayes’ theorem to calculate the post-test probability, using the following formula:

\[
\text{post-test probability} = \frac{\text{LR} \times (\text{pretest probability} / (1-\text{pre-test probability}))}{(\text{LR} \times (\text{pretest probability} / (1-\text{pre-test probability}))) + 1}
\]

To compare study results we used pre-test probabilities extracted from the literature: a probability of 10% for endometrial cancer and a probability of 40% for focal benign or (pre)malign endometrial disease.

Results

Study selection

A total of nine systematic reviews assessed (part of) the diagnostic pathway for women with PMB and met the criteria for inclusion (Figure 1). Of the selected systematic reviews, four articles assessed the use of TVS, one described the use of SIS, two assessed the use of outpatient endometrial sampling and two assessed the use of hysteroscopy in patients with abnormal uterine bleeding. Table I shows further details of these studies.
Figure 1. Study selection diagram

283 potentially relevant studies identified and screened for retrieval from electronic search

215 studies excluded*

68 studies retrieved for more detailed evaluation

59 studies excluded*
- 9 different subject
- 9 review, not systematic
- 22 cohort study
- 8 author reply
- 3 RCT
- 3 cost-effectiveness
- 3 withdrawn
- 2 case report
- 1 other

9 systematic reviews included, after consensus by two independent reviewers

*The reference list for excluded studies is available from the corresponding author.

Five of the selected reviews included both pre- and post-menopausal women.21-25 For this review, we used only the calculations and conclusions concerning postmenopausal women. The diagnostic accuracy reported in all included reviews is shown in Table 2. Additionally, we identified a set of five national and international guidelines concerning diagnostic strategies for PMB.7,10-13
### Table 1. Studies about accuracy of diagnostic tests in women with PMB

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Type of study</th>
<th>Studies (N)</th>
<th>Patients (N)</th>
<th>Subject</th>
<th>Failure/ inadequate</th>
<th>Reference test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total PMP</td>
<td>Total PMP</td>
<td>HRT</td>
<td>Overall PMP</td>
<td></td>
</tr>
<tr>
<td>Smith-Bindman et al.</td>
<td>1998</td>
<td>systematic review</td>
<td>35</td>
<td>35</td>
<td>5892</td>
<td>5892</td>
<td>8.0% TVS endometrial sampling or D&amp;C or hysterectomy</td>
</tr>
<tr>
<td>Tabor et al.</td>
<td>2002</td>
<td>systematic review</td>
<td>9</td>
<td>9</td>
<td>3813</td>
<td>2700</td>
<td>0.0% TVS D&amp;C or endometrial sampling</td>
</tr>
<tr>
<td>Gupta et al.</td>
<td>2002</td>
<td>systematic review</td>
<td>57</td>
<td>4</td>
<td>7270</td>
<td>464</td>
<td>1.3% TVS endometrial sampling or D&amp;C or hysterectomy</td>
</tr>
<tr>
<td>Timmermans et al.</td>
<td>2010</td>
<td>syst review IPD meta-analysis</td>
<td>11</td>
<td>11</td>
<td>2896</td>
<td>2896</td>
<td>n/a TVS n/a histology (different methods)</td>
</tr>
<tr>
<td>Clark et al.</td>
<td>2001</td>
<td>systematic review</td>
<td>6</td>
<td>2</td>
<td>881</td>
<td>220</td>
<td>4% biopsy 2% histology (by inpatient sampling)</td>
</tr>
<tr>
<td>Dijkhuizen et al.</td>
<td>2000</td>
<td>systematic review</td>
<td>39</td>
<td>7</td>
<td>7814</td>
<td>760</td>
<td>0.41% biopsy 0.54% D&amp;C or hysterscopy or hysterectomy</td>
</tr>
<tr>
<td>de Kroon et al.</td>
<td>2003</td>
<td>systematic review</td>
<td>24</td>
<td>n/a</td>
<td>2278</td>
<td>268</td>
<td>&lt;5% SIS 7% n/a hysterotomy</td>
</tr>
<tr>
<td>Clark et al.</td>
<td>2002</td>
<td>systematic review</td>
<td>65</td>
<td>2</td>
<td>26346</td>
<td>1948</td>
<td>4.3% hysteroscopy 3.6% histology (different methods)</td>
</tr>
<tr>
<td>van Dongen et al.</td>
<td>2007</td>
<td>systematic review</td>
<td>17</td>
<td>5</td>
<td>4208</td>
<td>306</td>
<td>4.4% hysteroscopy 3.2% histology (different methods)</td>
</tr>
</tbody>
</table>

n/a = not applicable
PMP = postmenopausal women
* pre- and postmenopausal women
Table 2. Diagnostic accuracy of diagnostic tests in women with PMB

<table>
<thead>
<tr>
<th>Study</th>
<th>Subject</th>
<th>Outcome measure</th>
<th>Cut-off value or device</th>
<th>Sens* (%)</th>
<th>Spec* (%)</th>
<th>LR**</th>
<th>LR-*</th>
<th>Pre-test probability*</th>
<th>Post-test probability*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith-Bindman</td>
<td>TVS</td>
<td>endometrial cancer</td>
<td>5 mm</td>
<td>96.0</td>
<td>61.0</td>
<td>0.07</td>
<td>10.0%</td>
<td>1.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial cancer</td>
<td>4 mm</td>
<td>96.0</td>
<td>53.0</td>
<td>0.08</td>
<td>10.0%</td>
<td>0.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial cancer</td>
<td>3 mm</td>
<td>100.0</td>
<td>38.0</td>
<td>0.00</td>
<td>10.0%</td>
<td>0.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial disease</td>
<td>5 mm</td>
<td>95.0</td>
<td>92.0</td>
<td>0.05</td>
<td>40.0%</td>
<td>3.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial disease</td>
<td>4 mm</td>
<td>91.0</td>
<td>69.0</td>
<td>0.13</td>
<td>40.0%</td>
<td>8.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial disease</td>
<td>3 mm</td>
<td>98.0</td>
<td>62.0</td>
<td>0.03</td>
<td>40.0%</td>
<td>2.11%</td>
<td></td>
</tr>
<tr>
<td>Tabor</td>
<td>TVS</td>
<td>endometrial cancer</td>
<td>MoM</td>
<td>96.0</td>
<td>50.0</td>
<td>0.08</td>
<td>10.0%</td>
<td>0.9%</td>
<td></td>
</tr>
<tr>
<td>Gupta ‡</td>
<td>TVS</td>
<td>endometrial cancer</td>
<td>5 mm</td>
<td>91.7</td>
<td>66.1</td>
<td>0.16</td>
<td>10.0%</td>
<td>2.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial cancer</td>
<td>4 mm</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial disease</td>
<td>5 mm</td>
<td>95.7</td>
<td>77.5</td>
<td>0.08</td>
<td>40.0%</td>
<td>5.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial disease</td>
<td>4 mm</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Timmermans</td>
<td>TVS</td>
<td>endometrial cancer</td>
<td>5 mm</td>
<td>90.3</td>
<td>54.0</td>
<td>0.18</td>
<td>10.0%</td>
<td>2.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial cancer</td>
<td>4 mm</td>
<td>94.8</td>
<td>46.7</td>
<td>0.11</td>
<td>10.0%</td>
<td>1.2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial cancer</td>
<td>3 mm</td>
<td>97.9</td>
<td>35.4</td>
<td>0.06</td>
<td>10.0%</td>
<td>0.7%</td>
<td></td>
</tr>
<tr>
<td>Clark</td>
<td>Sampling</td>
<td>endometrial cancer</td>
<td>all devices</td>
<td>91.9</td>
<td>99.7</td>
<td>95.1</td>
<td>4.5%</td>
<td>81.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>hyperplasia</td>
<td>all devices</td>
<td>81.0</td>
<td>98.9</td>
<td>73.6</td>
<td>14.3%</td>
<td>66.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pipette</td>
<td>Pipette</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>14.3%</td>
<td>62.3%</td>
<td></td>
</tr>
<tr>
<td>Dijkhuizen</td>
<td>Sampling</td>
<td>endometrial cancer</td>
<td>all devices</td>
<td>95.0</td>
<td>99.5</td>
<td>190.0</td>
<td>10.0%</td>
<td>95.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pipette</td>
<td>Pipette</td>
<td>99.6</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pipette</td>
<td>Pipette</td>
<td>88.0</td>
<td>98.0</td>
<td>44.0</td>
<td>14.3%</td>
<td>88.0%</td>
<td></td>
</tr>
<tr>
<td>de Kroon</td>
<td>Hydroscopy</td>
<td>hyperplasia</td>
<td>n/a</td>
<td>95.0</td>
<td>88.0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Clark</td>
<td>Hydroscopy</td>
<td>endometrial cancer</td>
<td>n/a</td>
<td>86.4</td>
<td>99.2</td>
<td>60.9</td>
<td>3.9%</td>
<td>71.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>endometrial cancer (PMP)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>38.3</td>
<td>10.0%</td>
<td>64.8%</td>
<td></td>
</tr>
<tr>
<td>van Dongen</td>
<td>Hydroscopy</td>
<td>abnormal cavity</td>
<td>n/a</td>
<td>96.0</td>
<td>90.0</td>
<td>7.9</td>
<td>61.0%</td>
<td>93.0%</td>
<td></td>
</tr>
</tbody>
</table>

n/a = not applicable or not available; MoM = multiples of the median; PMP = postmenopausal.

* Bold when extracted from article, normal when calculated, italic when extracted from literature.
†‡ For the cut-off value of 4 mm, no studies included with good quality.
¥ Pre- and postmenopausal women.
Quality assessment

The results of the quality assessment are reported in Figure 2. Overall, study quality was good. The quality of the formulation of the objective was rated as moderate (according to the checklist, studies scored well on this item if they described the patient population, the intervention, the reference standard and the desired result). All studies scored positively on items concerning the literature search, description of study characteristics and correctly performed meta-analysis. Study quality assessment was adequately performed in 55% (5/9) of the reviews.

Figure 2. Methodological quality of included studies

Transvaginal sonography

The main goal of TVS is to exclude endometrial cancer. Almost every guideline refers to a meta-analysis performed in 1998 by Smith-Bindman et al. It used traditional statistical methods to combine the data from 35 published studies regarding the use of TVS in the evaluation of women with PMB. Using the reported data from each study, 2 × 2 tables were constructed of endometrial thickness measured by TVS (above or below a threshold) against the presence or absence of endometrial cancer. Only 16 of the 35 included studies reported the number of women who could not tolerate TVS (mean 0%) and only 14 studies reported non-diagnostic results (mean 0%). With a cut-off value of 5 millimetres (mm), the sensitivity for detecting endometrial cancer was 96%, and the specificity 61%. This combination of sensitivity and specificity reduces a pre-test probability of 10% for endometrial cancer to a post-test probability (for a negative test) of 1%. Thus, based on the post-test probability of 1%, conservative management is recommended to women with an endometrial thickness of ≤5 mm.
The three other meta-analyses of TVS reached different conclusions, however. Tabor et al conducted a meta-analysis of nine studies. They included studies only if the corresponding author was able to supply original data. For each included study, the median endometrial thickness per centre was calculated and multiples of the median were used to pool data. They chose not to use a cut-off value, because there were statistically significant differences in endometrial thickness between centres, which may reflect differences in the populations studied or in the method of measuring endometrial thickness by TVS. In this study, a sensitivity of 96% and a specificity of 50% were found. These values give a post-test probability for a negative test of about 1% with a pre-test probability of carcinoma of 10%. These results are comparable to those of Smith-Bindman et al, but the authors disagreed on the interpretation of the results. The conclusion of Tabor et al was that a 4% false-negative rate is not acceptable and therefore the use of TVS in the evaluation of PMB is not recommended prior to invasive testing.

Gupta et al performed a systematic quantitative review in which they focused on study quality assessment. None of the nine studies that used a cut-off for endometrial thickness of ≤4 mm were of good quality. Only four studies (out of 21) used a ≤5 mm cut-off, but these employed the best quality criteria. Pooling of the results of these four studies resulted in a LR− of 0.16. This LR implies that a patient with a negative test result (endometrial thickness ≤5 mm) and pre-test probability of 10% would have a post-test probability of 2.5%. Their conclusion was that TVS can be used to rule out endometrial hyperplasia or carcinoma using an endometrial thickness of ≤5 mm.

In conclusion, the meta-analyses done by Tabor, Gupta and Smith-Bindman are limited because they are based on previously published data, and probably overestimate the accuracy of predictions based on endometrial thickness. With respect to meta-analysis of randomised trials, individual patient data are considered to be superior to meta-analysis of the literature. The use of individual patient data instead of published summary data gives less optimistic but more accurate conclusions. In diagnostic reviews the same might apply. Timmermans et al tried to overcome this limitation using a meta-analytic approach in which individual patient data from a series of original studies were combined. This study showed that in previous studies and meta-analyses, the diagnostic accuracy of TVS had been overestimated. Timmermans et al found a lower diagnostic accuracy for TVS than was reported previously: a sensitivity of 95% and a specificity of 47% at a cut-off of ≤4 mm, giving a post-test probability for a negative test of 1.2%. At a cut-off of ≤3 mm,
they found a sensitivity of 98% and a specificity of 35%, giving a LR for a negative test result of 0.06. Thus, a cut-off level of ≤3 mm reduced a pre-test probability of 10% to a 0.7% post-test probability. The conclusion of this article was that the use of TVS measurement of endometrial thickness remains justified, but with a recommended cut-off level of ≤3 mm.

**Outpatient endometrial sampling**

Clark et al conducted a systematic review and meta-analysis to determine the diagnostic accuracy of outpatient endometrial sampling in detecting endometrial hyperplasia. Postmenopausal women were included in two studies, in which they represented 25% of the (combined) patient sample. In these two articles three different diagnostic devices had been used for endometrial sampling: Accurette®, Pipelle® and Vabra® aspirator. The pre-test probability of 14.3% was increased to a post-test probability for a positive result of 66.7% (95% CI 42.3–83.9%). In 34 of 149 women the endometrial sampling was unsuccessful, with a failure rate (sampling not possible for technical reasons) of 17% (25/149) and an inadequate sampling rate (not enough tissue obtained for a pathologic diagnosis) of 7% (9/124). This review demonstrated that endometrial sampling is moderately accurate in diagnosing (pre) malignant endometrial pathology. A positive test result was more accurate than a negative test result (sensitivity 91.9%, with a specificity of 99.7%). Clark et al concluded that the more clinically significant the endometrial pathology is, the better the diagnostic accuracy of outpatient endometrial sampling will be and, hence, the more clinically useful the test. Additional endometrial assessment should be undertaken with technical failure or inadequate sampling, especially if symptoms persist.

Dijkhuizen et al performed a meta-analysis to assess the accuracy of endometrial sampling devices in the detection of endometrial cancer and atypical hyperplasia. With respect to the diagnosis of endometrial cancer, they identified seven studies that were limited to postmenopausal women. The pooled data from these studies showed a sensitivity of 95% and a specificity of 99.5%, giving a post-test probability for a positive test of 95.5%. Outpatient endometrial sampling therefore appeared to be a highly sensitive technique for diagnosing endometrial cancer. With regard to inadequate sampling (0 to 54% of cases in the studies they reviewed), they concluded that an inadequate sample is an indication for further investigation, based on an article of Farrell et al which demonstrated that of those women for whom the result of the Pipelle was ‘insufficient’ 20% had uterine pathology after further investigation, 3% with endometrial cancers.
Saline infusion sonography

Only one systematic review, by de Kroon et al has described the evaluation of the diagnostic accuracy of SIS in pre- and post- menopausal women with abnormal uterine bleeding. The main outcome measures were LRs, post-test probabilities and the success rate of SIS in the prediction of uterine cavity abnormality. We focus here on the results regarding women with PMB. The review identified 24 studies with homogeneous data, but only five of these concerned postmenopausal women. Pooling the data from these five articles gives a sensitivity of 95% and a specificity of 88%. The calculations for endometrial cancer were not mentioned. Sensitivity and specificity were not separately described for pre- and postmenopausal women, but the overall success rate was significantly lower in postmenopausal women: 87%, compared with 95% for premenopausal women (P < 0.01).

Hysteroscopy

Clark et al 2002 performed a systematic quantitative review in which they focused on the diagnostic accuracy of hysteroscopy in diagnosing endometrial cancer or hyperplasia. Postmenopausal women represented 29% of the populations studied. Only two studies concerning postmenopausal women were rated as high quality. Pooled data from these showed a post-test probability of a positive test of 71.8% (95% CI 67.0–76.6%). In these studies the failure rate for hysteroscopy (ambulant or inpatient) was 3.4% (67 of 1948 women), which was comparable with the overall failure rate in premenopausal women. Sensitivity and specificity were not separately described for pre- and postmenopausal women, but the overall sensitivity and specificity were 86.4% and 99.2% respectively. The authors concluded that when the uterine cavity is adequately visualized, hysteroscopy is highly accurate and clinically useful in the diagnosis of endometrial cancer. However, its high accuracy relates to diagnosing cancer rather than its exclusion.

Another systematic review and meta-analysis of diagnostic hysteroscopy was performed by van Dongen et al. This article focused on studies of the use of hysteroscopy in the diagnosis of intrauterine abnormalities, rather than endometrial cancer per se, because Clark et al had already shown in their meta-analysis that diagnostic hysteroscopy is accurate in the diagnosis of endometrial cancer. In this review five studies of postmenopausal women with homogeneous data were included. The pooled sensitivity and specificity in the assessment of uterine cavity abnormality were 96% (95% CI 93–99%) and 90% (95% CI 83–95%) respectively. With a pre-test probability of uterine cavity abnormalities of
61.0% (the prevalence in this group), they found a post-test probability for a positive test of 93% (95% CI 88–95%). The conclusion was that this meta-analysis gives strong evidence that diagnostic hysteroscopy is accurate in the diagnosis of intrauterine abnormalities.

**International guidelines**

The published national and international guidelines describe different diagnostic pathways in the diagnostic work-up of women with postmenopausal bleeding. When a patient presents with PMB the first step in every guideline is referral to a gynaecologic practice for examination, pap smear and TVS. Only the US guidelines recommend either TVS or outpatient endometrial sampling as the first step in diagnosing women with PMB, based on similar sensitivities and cost-effectiveness for the detection of endometrial cancer for an endometrial thickness of 5 mm or more and for endometrial sampling when ‘sufficient’ tissue is obtained. In the other guidelines the first step is TVS, based on the high sensitivity and non-invasive character of the procedure. Different guidelines use different cut-off values of endometrial thickness, varying from 3 to 5mm. These cut-off points are mostly based on the meta-analysis by Smith-Bindman, but also on Swedish literature, and the review by Gupta et al. The most important issue is what probability of endometrial cancer is deemed acceptable after a negative test.

In the US guidelines, endometrial sampling is recommended with a cut-off value for the endometrial thickness of 5 mm and at the same time they recommend TVS when the endometrial sampling is deemed ‘insufficient’. SIS is used to distinguish between a diffusely thickened endometrium, for which D&C could be the next step, and between a focal lesion, for which a hysteroscopy is the next advised step. The National Guideline Clearinghouse stated that D&C in women with PMB should be performed only when endometrial sampling is indicated and cannot be performed or is inconclusive and sonographic techniques are non-reassuring. D&C should always have concomitant hysteroscopy, in case of focal pathology.

The European guidelines advise endometrial sampling only when the endometrial thickness is above the cut-off value, possibly together with a SIS to distinguish between diffuse and focal pathology. With focal lesions the recommendation is to perform a (therapeutic) hysteroscopy and with diffuse lesions D&C, but only when endometrial sampling is insufficient or has failed. Where the endometrium is thin, the guidelines recommend conservative management. Only the Scottish guideline recommends further investigation if the clinician, the patient or both are not
reassured. The exact sequence of investigation will depend upon clinical judgment, local resources, local expertise and patient preference. With recurrent or persistent PMB there are different strategies. In the Dutch guideline immediate hysteroscopy is advised; the Swedish guideline advises outpatient endometrial sampling or, if technically not possible, D&C. The further investigation and management of benign lesions in women with PMB require more research. The question is whether the treatment of benign lesions improves the patient’s quality of life, morbidity and survival.

Discussion

The goal of this systematic review was to produce an evidence-based diagnostic pathway for patients with PMB. The most important conclusion is that in neither systematic reviews nor international guidelines can consensus be found regarding the sequence in which the different procedures should be implemented. All four types of test have been shown to be accurate and feasible in excluding or diagnosing endometrial cancer; by their high sensitivities and specificities.

Based on the available evidence, discussed in this review, we can conclude that TVS is an accurate method to exclude endometrial cancer, although there is still debate over the best cut-off value for endometrial thickness that should warrant endometrial sampling. Based on the highest sensitivity, a cut-off value of 3 mm is recommended, but the cost-effectiveness of this value has yet to be demonstrated.

Regarding the use of outpatient endometrial sampling, the two included reviews showed the high accuracy of this diagnostic method. Clark et al focused mainly on hyperplasia, while Dijkhuizen et al concluded that outpatient endometrial sampling is an accurate method for excluding cancer in women with PMB. However, the technique had a high rate of insufficient or failed sampling (0–54% in different observational studies). An insufficient sample should be an indication for further investigation. From the available reviews we cannot draw conclusions regarding the sequence of the tests in the diagnostic pathway of PMB. All the studies evaluated the tests independently, without any consideration of combinations of tests or previous test results.

In determining the best sequence of tests, different factors have to be taken into account, such as cost-effectiveness, the prevalence of endometrial cancer, local logistics (the availability of ultrasound, the use of outpatient endometrial sampling and
the use of outpatient hysteroscopy), as well as doctor and patient preferences. The preferences of doctors in relation to diagnostic procedures for endometrial cancer in women with PMB have not been investigated. Furthermore, guidelines need to meet the expectations of the patients; most women want to rule out endometrial cancer with a certainty of 100% and they are prepared to undergo rather invasive and painful diagnostic tests in order to achieve this. However, a post-test probability of 0.0% seems virtually impossible and one should also keep in mind that the risk of endometrial cancer in a population of asymptomatic postmenopausal women is reported to be 0.2%.

Clark et al determined the most cost-effective strategy for diagnosing endometrial cancer. They constructed a decision model and evaluated 12 different strategies for the initial investigation of PMB. With a cancer probability of 10%, the strategy with TVS as the initial test with a cut-off of 4 mm followed by endometrial sampling was most cost-effective. Unfortunately, a cut-off value of 3 mm was not considered in their evaluation. More importantly, in this decision model, the assumptions made regarding test accuracy were based on the available systematic reviews. Systematic literature reviews in diagnostic research report the accuracy of tests, and thereby assist clinicians in their decision-making. However, there are limitations to this approach, as the analysis of such data often does not allow reviewers to explore the diagnostic information gained from combinations of tests. In clinical practice, tests are commonly combined in diagnostic sequences and disease probabilities are usually estimated in a hierarchical manner, first combining information from the history and examination, followed by additional information obtained from other diagnostic procedures (e.g., TVS, endometrial sampling). Studies of test accuracy often do not take this clinical paradigm into account, but tend to report test results in isolation and disregard the history and examination. In addition, they usually analyse a single test at a time, without taking into account of what is known from previous testing.

There is considerable variability in endometrial thickness and the likelihood of endometrial cancer across women. Individual patient characteristics, including age, time since menopause, obesity, hypertension, diabetes mellitus and reproductive factors, are associated with a higher prevalence of endometrial cancer. However, current policy is not based on these risk factors, but only on endometrial thickness. Breijer et al developed an algorithm for diagnostic pathways in women with PMB. This algorithm includes the calculation of the pre-test probability of endometrial cancer based on individual patient characteristics and the diagnostic approach to benign pathology, both of which require further research.
Chapter 2

**Figure 3.** Possible diagnostic pathway in postmenopausal bleeding

The dashed box shows that SIS can be used to distinguish between focal and diffuse pathology before performing a hysteroscopy, according to the local protocol.

Future research should also aim to maximise accuracy in relation to cost-effectiveness for the different methods. Incorporation of a combination of endometrial thickness and patient characteristics within a single diagnostic pathway increased diagnostic accuracy in some studies.\(^{40-42}\) Future research has to focus on the combination of different diagnostic tests as well as the incorporation of patient characteristics, rather than on the diagnostic accuracy of a single test. Furthermore, by combining and analysing individual patient data from different studies (i.e. meta-analysis of individual patient data), larger databases can be obtained, in which previously described models can be externally validated.\(^{40-43}\)
We can conclude that the first step in the diagnostic pathway should be the measurement of endometrial thickness, using a cut-off point of 3 or 4 mm, followed by endometrial sampling. Figure 3 shows an algorithm with an evidence-based diagnostic pathway for women with PMB. Only when TVS is not readily available should direct endometrial sampling be an option. For further investigation when the sample is insufficient or when it is unsuccessful, SIS can be used to distinguish between focal and diffuse pathology. Hysteroscopy should be used as the final step in the diagnostic pathway of women with PMB.

Conclusions

• Neither in systematic reviews nor in international guidelines is consensus found regarding the best sequence of diagnostic procedures for women with PMB.

• Measurement of endometrial thickness, endometrial sampling and hysteroscopy have been independently shown to be accurate in excluding endometrial cancer.

• In relation to endometrial thickness, a cut-off value of 3 or 4 mm is recommended, but the cost-effectiveness of this strategy has yet to be demonstrated.
Appendix A: Search strategy - Medline

# Searches
1  Postmenopause/
2  postmenopau*.tw.
3  post-menopau*.tw.
4  or/1-3
5  exp Hemorrhage/
6  (bleed* or hemorrhag* or haemorrhag* or blood loss*).tw.
7  or/5-6
8  4 and 7
9  hysteroscopy/
10 hysteroscop*.tw.
11 or/9-10
12 Biopsy/ or Biopsy, Needle/
13 exp Curettage/
14 (biop* or curett* or pipelle*).tw.
15 or/12-14
16 ultrasonography/ or endosonography/ or exp ultrasonography, doppler/
17 vagina/us or endometrium/us or Uterine Hemorrhage/us or exp Uterine Neoplasms/us or exp Uterine Diseases/us
18 ((endometr* or vagina* or endovagin* or transvagina* or trans-vagina* or endo-vagi* or uter* or intrauter* or intra-uteri*) adj6 (echo* or ultrasound or ultrasono* or sonograph* or doppler or endoscop* or endoson*)).tw.
19 (hysterosalpingogr* or hysterosonogra*).tw.
20 thick*.tw.
21 or/16-20
22 11 or 15 or 21
23 8 and 22
24 exp "sensitivity and specificity"/
25 (diagnos* or test or tests or exclude or value or role or evaluation).ti.
26 (accurac* or (sensitivit* and specificit*) or (predictive adj3 value*1) or (false adj2 (negative or positive)) or ROC or pretest or validat*).tw.
Diagnostic evaluation

or/24-26
11 or 15 or 21 or 27
28 and 8
(meta-analysis.pt. or exp technology assessment, biomedical/ or ((hta or health technology) adj6 assessment*) or meta analy* or metaanaly* or metaanaly*).tw. or (cochrane or evidence or EBM).jw. or ((review* or search*) adj10 (literature* or medical database* or medline or pubmed or embase or cochrane or cinahl or psychinfo or psychlit or healthstar or biosis or current conten* or systemat*)).tw.) not (comment or editorial or historical-article).pt.
29 and 30
29
limit 32 to yr="2008 -Current"
33 not 31
review.pt.
35 (review or overview).ti.
36
37
38
34 and 37
39
40 ("11042572" or "19576369" or "17516956" or "12039131" or "9809732" or "12350192" or "12225294" or "14550365").an.
41
40 and 31
42
40 and 38
from 31 keep 1-59
from 38 keep 1-4
from 39 keep 1-115

Appendix B: Search strategy - Embase

# Searches
1 Postmenopause/
2 (postmenopau* or post-menopau*).tw.
3 (after menopaus* or after the menopaus* or following menopaus* or following the menopaus*).tw.
4 (older adj2 (wom#n or female*)).tw.
5 or/1-4
Chapter 2

6 exp bleeding/
7 (bleed* or hemorrhag* or haemorrhag* or blood loss*).tw.
8 or/6-7
9 5 and 8
10 hysteroscopy/
11 hysteroscop*.tw.
12 or/10-11
13 biopsy.mp.
14 curettage/
15 (biop* or curett* or pipelle*).tw.
16 or/13-15
17 exp echography/
18 hysteroscopy/ or hysterosalpingography/ or hysterography/
19 thickness/
20 ((endometr* or vagina* or endovagin* or transvagina* or trans-vagina* or endo-vagl* or uter* or intrauter* or intra-uteri*) adj6 (echo* or ultrasound or ultrasono* or sonograph* or doppler or endoscop* or endoson*)).tw.
21 (hysterosalpingogr* or hysterosonogra*).tw.
22 thick*.tw.
23 or/17-22
24 diagnostic accuracy/ or diagnostic test/ or diagnostic value/ or exp diagnostic error/ or roc curve/ or “sensitivity and specificity”/ or validity/ or predictive validity/
25 (diagnos* or test or tests or exclude or value or role or evaluation).ti.
26 (accurac* or (sensitivit* and specificit*) or (predictive adj3 value*1) or (false adj2 (negative or positive)) or ROC or pretest or validat*).tw.
27 or/24-26
28 12 or 16 or 23 or 27
29 28 and 9
30 uterus bleeding/di or vaginal bleeding/di
31 30 and 5
32 29 or 31
Diagnostic evaluation

33 exp meta analysis/ or exp Literature/ or exp Biomedical Technology Assessment/ or (hta or (health technology adj6 assessment*) or metaanaly* or meta analy* or meta?analy*).tw. or (cochrane or evidence or EBM).jx. or ((review* or search*) adj10 (literature* or medical data base* or medline or pubmed or embase or cochrane or cinahl or psychinfo or psychlit or healthstar or biosis or current content* or systematic*)).tw.

34 32 and 33
35 limit 32 to yr="2008 -Current"
36 35 not 34
37 review.pt.
38 (review or overview).ti.
39 37 or 38
40 36 and 39
41 36 not 39
42 (“2009314425” or “2007250425” or “2003399129” or “2002343827” or “2002331806” or “2002111470” or “2000376745” or “1998375444”).an.
43 42 and 34
Chapter 2

References


42. Opmeer BC, van Doorn HC, Heintz AP, Burger CW, Bossuyt PM, Mol BW. Improving the existing diagnostic strategy by accounting for the characteristics of the women in the diagnostic work-up in the postmenopausal bleeding. BJOG 2007 Jan;114(1):51-58.