Is this woman perimenopausal?
Bastian L A, Smith C M, Nanda K

CRD summary
This review assessed the accuracy of signs, symptoms and laboratory tests to determine whether a woman is perimenopausal. It found that no one symptom or test is accurate enough by itself to rule in or rule out perimenopause. Although the review suffered from some methodological weaknesses, the results do support these conclusions.

Authors' objectives
To review the accuracy of self-assessment, symptoms, signs and laboratory tests in diagnosing women in perimenopause.

Searching
MEDLINE was searched from 1966 to 2001; the search terms were reported. The search was restricted to studies reported in the English language. The reference lists of included studies and other publications were screened for additional studies.

Study selection
Study designs of evaluations included in the review
Studies that included a premenopausal control group were eligible for inclusion. Studies that included a young control group (i.e. 20-year-old women) or an older postmenopausal group (i.e. 60- to 70-year-old women) were excluded. However, the study designs actually included in the review were cross-sectional or prospective studies.

Specific interventions included in the review
Studies that assessed laboratory tests were eligible for inclusion. The laboratory tests assessed in the included studies measured inhibins and follicle-stimulating hormone (FSH). The included studies assessed different combinations of the following signs and symptoms: hot flash, mood, insomnia, nervous tension, sexual interest, night sweat, incontinence, vaginal dryness, psychologic distress and cigarette smoking.

Reference standard test against which the new test was compared
Studies that used the diagnosis of perimenopause based on the menstrual irregularity of 3 to 11 months of amenorrhoea were eligible for inclusion.

Participants included in the review
Studies that included men or that focused on cancer or osteoporosis were excluded. Studies that included women receiving hormone replacement therapy were also excluded. The women included in the studies were aged from 40 to 60 years and were of Thai, Australian, American, Chinese, British and Dutch origin. The proportion of perimenopausal women in each study ranged from 7 to 70%.

Outcomes assessed in the review
Studies that reported sufficient data to calculate the sensitivity and specificity were eligible for inclusion.

How were decisions on the relevance of primary studies made?
Two authors systematically reviewed and identified titles and abstracts for content and quality.

Assessment of study quality
The studies were given a grade of A, B or C based on the study design and level of evidence. Grade A required an independent blind comparison of signs or symptoms with a 'gold' standard of diagnosis among a large number of consecutive patients suspected of having the target condition. Grade B required an independent, blind comparison of
signs or symptoms with a ‘gold’ standard of diagnosis among a small number of consecutive patients suspected of having
the target condition. Grade C included all other studies. Two authors assessed validity independently using a
standardised abstraction form. A third author resolved any discrepancies about quality.

Data extraction
Two authors extracted the data using a standardised abstraction form (it was unclear whether this was carried out
independently). The sensitivity, specificity, and positive and negative likelihood ratios (LRs), together with their 95% confidence intervals (CIs), were calculated.

Methods of synthesis
How were the studies combined?
In the absence of heterogeneity (P<0.05), pooled estimates of sensitivity, specificity, and positive and negative LRs, together with their 95% CIs, were calculated using random-effects models. For heterogeneous data, ranges in values were reported.

How were differences between studies investigated?
Estimates of sensitivity, specificity, and positive and negative LRs were assessed for heterogeneity using the chi-squared statistic.

Results of the review
Sixteen studies were included in the review. These studies involved a total of 42,818 women, of whom 22,321 were premenopausal and 20,497 were perimenopausal.

Five studies were evidence level A, the other 11 were level B.

Hot flashes (n=5): the sensitivity ranged from 22 to 55% and the specificity from 83 to 91%. The positive LR ranged from 2.15 to 4.06 and the negative LR from 0.54 to 0.87.

Night sweats (n=3): the sensitivity ranged from 20 to 50% and the specificity from 74 to 87%. The positive LR ranged from 1.48 to 3.79 and the pooled negative LR was 0.92 (95% CI: 0.91, 0.93).

Vaginal dryness (n=3): the sensitivity ranged from 11 to 29% and the specificity from 80 to 97%. The positive LR ranged from 1.09 to 3.79 and the pooled negative LR was 0.92 (95% CI: 0.91, 0.04).

Incontinence (n=4): the sensitivity ranged from 16 to 39% and the specificity from 64 to 91%. The positive LR ranged from 1.30 to 3.14 and the negative LR from 0.82 to 0.94.

Depressed mood (n=5): the sensitivity ranged from 9 to 47% and the specificity from 64 to 97%. The positive LR ranged from 0.30 to 3.14 and the negative LR from 0.82 to 0.94.

Insomnia (n=5): the sensitivity ranged from 21 to 53% and the specificity from 63 to 83%. The positive LR ranged from 0.98 to 2.06 and the negative LR from 0.79 to 1.0.

Nervous tension and/or irritability (n=3): the sensitivity ranged from 41 to 59% and the specificity from 51 to 68%. The pooled positive LR was 1.23 (95% CI: 1.12, 1.34) and the pooled negative LR was 0.83 (95% CI: 0.77, 0.90).

Self-rating (n=2): the sensitivity ranged from 77 to 94% and the specificity from 39 to 64%. The positive LR ranged from 1.53 to 2.13 and the negative LR from 0.18 to 0.36.

The following symptoms or tests were each assessed in one study: psychologic distress, sexual interest, current-smoking, FSH levels, inhibin A and B levels, and IR-INH levels. The positive LRs ranged from 1.31 to 3.06 and the negative LRs from 0.45 to 0.97.
Authors' conclusions
No single element of the history or clinical examination is powerful enough to confirm the probability of being perimenopausal. Besides menstrual history, the most powerful predictor of menopausal status may be age.

CRD commentary
The review addressed a clearly defined objective. Although inclusion criteria were reported, these were repeated and slightly confusing. The literature search was limited to one database (MEDLINE) combined with screening references of included studies. No attempts were made to identify unpublished studies and the review was limited to English language studies. It is therefore possible that important studies might have been missed and that the review might be subject to publication bias. Some details of the review process were reported and all stages involved two reviewers. A formal quality assessment was undertaken; however, this involved stratifying studies into levels of evidence and the only criterion that really contributed to this was study size (all studies were graded A or B, these grades only differed according to whether a large or small sample was included). An assessment of individual quality items would have been more informative.

Adequate details of the studies, including the results of the quality grading, were tabulated. It would have been preferable to have also included full diagnostic accuracy results for the individual studies, rather than just the ranges and pooled values reported in the summary tables. A simple pooling of sensitivity, specificity and LRs was used to pool the studies, although this appears appropriate given the data. A more informative analysis would have looked at whether various combinations of signs or symptoms would have been more accurate in diagnosing menopausal status, rather than looking at each sign and symptom individually. The authors’ conclusions regarding the signs and symptoms assessed are supported by the results presented, but should be interpreted with some degree of caution due to the limitations highlighted above. The conclusions also included a comment on the ability of age to predict the menopause, but this did not appear to have been assessed in the included studies and was not discussed in the 'Results' section of the review.

Implications of the review for practice and research
Practice: The authors stated that current evidence suggests that FSH measurement does not help the clinician to make a diagnosis of menopausal status.

Research: The authors stated that further research needs to be conducted to document the additional benefit of hormone level tests in making a diagnosis of perimenopause.

Bibliographic details

PubMedID
12588275

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

Indexing Status
Subject indexing assigned by NLM

MeSH
Adult; Climacteric /physiology; Contraceptives, Oral; Diagnosis, Differential; Female; Follicle Stimulating Hormone /blood; Humans; Hysterectomy; Medical History Taking; Middle Aged; Physical Examination; Smoking

AccessionNumber
12003008101

Date bibliographic record published
30/06/2006

**Date abstract record published**
30/06/2006

**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.