Antral follicle count in the prediction of poor ovarian response and pregnancy after in vitro fertilization: a meta-analysis and comparison with basal follicle-stimulating hormone level


CRD summary
The review assessed the predictive performance of antral follicle count as a test for ovarian reserve in in vitro fertilisation. The authors concluded that transvaginal ultrasonography is a noninvasive and easy-to-perform method that provides essential predictive information on ovarian responsiveness. The search was limited and the conclusion appears slightly overstated in comparison with the reported results.

Authors' objectives
To evaluate the predictive performance of antral follicle count (AFC) as a test for ovarian reserve in in vitro fertilisation (IVF) and in comparison with basal follicle-stimulating hormone (FSH) levels. Only the predictive performance of AFC will be reported in this abstract as the search strategy used for the comparison with basal FSH does not meet the inclusion criteria for DARE.

Searching
The authors searched MEDLINE (from January 1996 to October 2003) and checked the bibliographies of included studies.

Study selection
Study designs of evaluations included in the review
The studies had to report data on AFC and IVF outcomes to be eligible for inclusion. All of the included studies were diagnostic cohort studies.

Specific interventions included in the review
Studies that evaluated AFC under transvaginal ultrasonography were eligible for inclusion. Where reported, the ultrasonographic probes used in the included studies ranged from 4 to 7.5 MHz.

Reference standard test against which the new test was compared
The reference standard used was the occurrence of poor ovarian response and/or nonpregnancy. Poor response encompassed cycle cancellation, insufficient follicular growth or oocyte yields, according to the definition in the individual studies.

Participants included in the review
Studies of patients undergoing IVF treatment were eligible for inclusion. Most of the included studies reported on the outcome of one cycle per couple; others reported on the outcome of multiple outcomes.

Outcomes assessed in the review
The studies had to 2x2 data on AFC and ovarian response to be eligible for inclusion. Some studies used the data of more than one IVF cycle per couple.

How were decisions on the relevance of primary studies made?
One reviewer screened the abstracts to determine inclusion.

Assessment of study quality
The studies were assessed in relation to sampling (consecutive versus other), data collection (prospective versus retrospective), study design (cohort versus case-control study), blinding, selection bias and verification bias. The authors did not state how many reviewers performed the validity assessment.
Data extraction
Two reviewers independently extracted the data. Any discrepancies were resolved by a third reviewer. Data were extracted to calculate sensitivity and specificity from published data or on the basis of information provided by the authors. Data on all cut-off points were extracted, and the number of patients was divided by the number of cut-off points to correct the weighting of the studies. The pre- and post-test probabilities were calculated for the outcomes, in addition to the proportion of patients or cycles with abnormal AFC.

Methods of synthesis
How were the studies combined?
Sensitivity-specificity points were plotted in receiver operating characteristic (ROC) space and a summary ROC curve was calculated. Study results were also tabulated.

How were differences between studies investigated?
The homogeneity of sensitivity and specificity were statistically assessed. The effects of study quality characteristics on the predictive performance were analysed in a logistic regression analysis. If a characteristic was found to have a significant impact a further analysis was undertaken. To investigate cut-off effects, sensitivities and specificities were correlated.

Results of the review
Eleven studies (1,760 IVF cycles) met the inclusion criteria.

For the prediction of poor ovarian response, the sensitivity ranged from 95% (corresponding specificity 69%) to 9% (corresponding specificity 97%) and the specificity from 97% (corresponding sensitivity 9%) to 40% (corresponding sensitivity 80%). The positive likelihood values ranged from 1.3 to 19.7. There was evidence of statistical heterogeneity (P<0.001).

For the prediction of nonpregnancy, the sensitivity ranged from 54% (corresponding specificity 87%) to 7% (corresponding specificity 83%) and the specificity from 98% (corresponding sensitivity 12%) to 33% (corresponding sensitivity 60%). The positive likelihood values ranged from 0.8 to 8.3. There was evidence of statistical heterogeneity.

The analysis of study characteristics did not explain differences in the predictive performance of the AFC test in the individual studies.

Authors’ conclusions
Transvaginal ultrasonography is a noninvasive and easy-to-perform method that provides essential predictive information on ovarian responsiveness.

CRD commentary
The review had clear inclusion criteria. The search for individual studies was limited, thus it is possible that potentially eligible studies might have been overlooked. Better documentation of the review process might enable a more thorough assessment of the potential for reviewer bias and errors. The details provided of the included studies highlighted the great variation in the results of the individual studies. Other than the graph no information about the summary ROC curve was given, so the appropriateness and fit of the model is difficult to evaluate. The authors’ conclusion appears slightly overstated given that only one study result showed a large positive likelihood ratio, the sensitivities and specificities varied greatly between studies, and the quality of the performance could not be predicted in the sensitivity analyses. Furthermore, it is not possible to comment on the comparison with basal FSH levels as this did not meet the criteria for inclusion on DARE.

Implications of the review for practice and research
Practice: The authors stated that AFC might be considered the test of first choice in the assessment of ovarian reserve prior to IVF.

Research: The authors did not state any implications for research.

**Bibliographic details**


**PubMedID**

15705365

**DOI**


**Other publications of related interest**


**Indexing Status**

Subject indexing assigned by NLM

**MeSH**

Female; Fertilization in Vitro; Follicle Stimulating Hormone /blood; Humans; Ovarian Follicle /cytology /ultrasonography; Predictive Value of Tests; Pregnancy; Pregnancy Outcome

**AccessionNumber**

12005009639

**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.