Systematic review of effectiveness of teaching critical appraisal

Hyde C, Parkes J, Deeks J, Milne R

Authors' objectives
To investigate the effect of teaching critical appraisal to health care professionals and consumers on a variety of outcomes related to their practice of health care and use of health care facilities and interventions.

Searching
MEDLINE, EMBASE, the Cochrane Library, PsycLIT, LISA, ERIC, CINAHL, SIGLE, the Social Sciences Citation Index and the Science Citation Index were searched up until December 1997; the search terms were reported. In addition, bibliographies of retrieved articles and reviews were checked and Internet searches were conducted. Medical education centres in the UK, Europe and USA were contacted, as were experts in the field. Only full articles were included in the review.

Study selection

Study designs of evaluations included in the review
Any comparative study was eligible for inclusion.

Specific interventions included in the review
Studies evaluating the effect of critical appraisal teaching were eligible for inclusion. The interventions included tutorials, study days, workshops, lectures, seminars, journal clubs and slide/tape programmes, alone or in combination; these were compared with no training in critical appraisal. The duration of these interventions ranged from 30 minutes to 16 hours.

Participants included in the review
Studies of participants in any health care setting, including health care students, professionals, managers, purchasers and health care users, were eligible for inclusion. Most of the studies were conducted in the USA on doctors or medical students.

Outcomes assessed in the review
The studies had to include at least one patient outcome (mortality, morbidity, quality of life or satisfaction), learner outcome (behaviour, critical appraisal skills, knowledge, attitudes or satisfaction), or teacher outcome (satisfaction) to be eligible for inclusion. The outcomes assessed in the included studies were skills, knowledge, attitudes and the behaviour of health professionals. There were no useable data on patient outcomes.

How were decisions on the relevance of primary studies made?
One reviewer initially screened the results of the searches. Studies potentially relevant to the review were retrieved, and three reviewers independently assessed them for inclusion using a proforma. Any disagreements were either resolved by discussion, or with the opinion of a fourth reviewer.

Assessment of study quality
Between-group comparison studies were assessed on the basis of: the method of randomisation and allocation concealment; blinding; similarity at baseline, or adjustment for differences at baseline; concurrent training interventions and potential contamination of the control group with critical appraisal teaching targeted at the experimental group; consistency of measurement between the intervention and control groups; the validity of the outcome being assessed; loss to follow-up; and the statistical analysis used (including analysis of a background trend, or learning effect for repeated use of an outcome measure).

Within-group comparisons were assessed on the basis of: similarity of the groups and pairing of responses; concurrent training interventions; consistency of measurement tool throughout the trial; the validity of the outcome measure;
blinding; loss to follow-up; and the statistical analysis used (including analysis of a background trend, or learning effect for repeated use of an outcome measure), and the minimisation of other biases.

One reviewer assessed quality, independently of the data extraction; while a second reviewer checked a sample (10%) for accuracy. A statistician checked statistical criteria.

**Data extraction**

One reviewer extracted the data using a standard form and a second reviewer checked the extraction. Any differences were resolved by consensus. A third reviewer checked a sample (10%) for accuracy.

**Methods of synthesis**

*How were the studies combined?*

The results were combined in a narrative using structured tabulation of the study characteristics, threats to validity and results, grouped by outcome measure. In addition, a limited meta-analysis was undertaken. The weighted mean differences and 95% confidence intervals for the assessment of knowledge through multiple-choice tests were calculated using a random-effects model.

*How were differences between studies investigated?*

A sensitivity analysis was undertaken to determine the effect of quality on the results.

**Results of the review**

Sixteen studies (n=2,997) were included in the review: one randomised controlled trial (RCT; n=44), 8 non-randomised controlled trials (n=808) and 7 before-and-after trials (n=2,145). The RCT and 6 of the controlled trials also assessed before and after effects.

**Skills.**

Eleven studies (16 comparisons) assessed skills. Twelve comparisons resulted in significant improvements in those receiving critical appraisal training, and one a significant improvement in the control group. The outcomes where improvement after critical appraisal training was reported included: diagnosis exercises; therapy exercises; concordance with expert appraisal; multiple-choice tests; study design, measurements and statistics used; sources of bias; and interpreting results, understanding their clinical significance, and using them in practice. The study favouring the control group assessed the appraisal of a test article. One study measuring perception of own skill showed no effect of critical appraisal training. Of the five comparisons from higher quality trials, four favoured critical appraisal training and one favoured the control.

**Knowledge.**

Ten studies (12 comparisons) assessed knowledge. Ten comparisons resulted in a significant improvement in those receiving critical appraisal training. Two of these assessments were by multiple-choice tests, two T/F test, one multiple-choice test and T/F test, three perception of own knowledge, and two used test questions. The outcomes in which improvements were seen included: concepts of literature evaluation, clinical epidemiological principles, research or study design, epidemiology, biostatistics, principles of critical appraisal, and understanding key terms in systematic reviews.

**Attitudes.**

Four studies (4 comparisons) assessed attitudes. Two of these reported significant improvement in those receiving critical appraisal training, with both assessing the agreement or disagreement with statements. Only one study was of high quality; this was one of those reporting a significant improvement in the agreement or disagreement with statements in those receiving critical appraisal training.

**Behaviour.**
Of 8 comparisons measuring 15 outcomes, only two (both by the same author) showed significant improvement after critical appraisal training; each measured one outcome. One of these studies reported an increased frequency in reading research articles, and the other a greater influence of these articles in reviewing clinical practice. Both of these studies were of poorer quality and were excluded from the sensitivity analysis.

Authors' conclusions
The authors concluded that teaching critical appraisal had positive effects on participants' attitudes, knowledge and skills. However, there were gaps in the evidence as to whether it impacts on decision-making, or patient health or satisfaction, and whether the benefit is of practical significance.

CRD commentary
The research question and the inclusion criteria were clearly stated. An extensive search was undertaken, with attempts to locate published and unpublished data. The initial screening of the search results and the validity assessment were both conducted by single reviewers, with only 10% being checked by a second reviewer; therefore, the potential for error and bias was increased during these stages of the review. The second stage of the study selection and data extraction processes were performed in duplicate, thus minimising the potential for bias and error. The grouping of the studies in the synthesis was appropriate and sufficient details of included studies were provided. The validity assessment was detailed and appropriate. On the whole, this was a well-conducted review, and the conclusions drawn by the authors seem to follow the data presented.

Implications of the review for practice and research
Practice: The authors stated that investing in critical appraisal teaching was likely to have had a positive impact. However, they stated that there was insufficient evidence to advise further expansion of critical appraisal activities.

Research: The authors advised further studies in partnership with adult educationalists and health care researchers. They also advised a large, multicentre, randomised trial evaluating the benefit of critical appraisal to postgraduates and continuing professional development, which assesses outcomes and changes significant to practice.

Funding
UK National R&D Programme, project number 12-8.

Bibliographic details

Original Paper URL
http://www.arif.bham.ac.uk/pdfs/TeachCritApp.PDF

Other publications of related interest

Indexing Status
Subject indexing assigned by CRD

MeSH
Biomedical Research /standards; Clinical Medicine /standards; Evidence-Based Medicine; Research Design /standards

AccessionNumber
12001008250

**Date bibliographic record published**
31/05/2005

**Date abstract record published**
31/05/2005

**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.