



Cochrane
Library

Cochrane Database of Systematic Reviews

Teaching critical appraisal skills in healthcare settings (Review)

Horsley T, Hyde C, Santesso N, Parkes J, Milne R, Stewart R

Horsley T, Hyde C, Santesso N, Parkes J, Milne R, Stewart R.
Teaching critical appraisal skills in healthcare settings.
Cochrane Database of Systematic Reviews 2011, Issue 11. Art. No.: CD001270.
DOI: [10.1002/14651858.CD001270.pub2](https://doi.org/10.1002/14651858.CD001270.pub2).

www.cochranelibrary.com

Teaching critical appraisal skills in healthcare settings (Review)
Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	3
BACKGROUND	4
OBJECTIVES	4
METHODS	4
RESULTS	6
Figure 1.	7
DISCUSSION	9
AUTHORS' CONCLUSIONS	10
ACKNOWLEDGEMENTS	10
REFERENCES	12
CHARACTERISTICS OF STUDIES	15
ADDITIONAL TABLES	24
APPENDICES	27
WHAT'S NEW	27
HISTORY	27
CONTRIBUTIONS OF AUTHORS	28
DECLARATIONS OF INTEREST	28
SOURCES OF SUPPORT	28
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	28
INDEX TERMS	28

[Intervention Review]

Teaching critical appraisal skills in healthcare settings

Tanya Horsley¹, Chris Hyde², Nancy Santesso³, Julie Parkes⁴, Ruairidh Milne⁵, Ruth Stewart⁶

¹Centre for Learning in Practice, Royal College of Physicians and Surgeons of Canada, Ottawa, Canada. ²Peninsula Technology Assessment Group (PenTAG), Peninsula College of Medicine & Dentistry, Exeter, UK. ³Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada. ⁴Public Health Medicine, Wessex Institute Health Research & Development, Southampton, UK. ⁵NETSCC, University of Southampton, Southampton, UK. ⁶Social Science Research Unit, Institute of Education, University of London, London, UK

Contact address: Tanya Horsley, Centre for Learning in Practice, Royal College of Physicians and Surgeons of Canada, 774 Echo Drive, Ottawa, Ontario, K1S5N8, Canada. thorsley@rcpsc.edu.

Editorial group: Cochrane Effective Practice and Organisation of Care Group.

Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 11, 2011.

Citation: Horsley T, Hyde C, Santesso N, Parkes J, Milne R, Stewart R. Teaching critical appraisal skills in healthcare settings. *Cochrane Database of Systematic Reviews* 2011, Issue 11. Art. No.: CD001270. DOI: [10.1002/14651858.CD001270.pub2](https://doi.org/10.1002/14651858.CD001270.pub2).

Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Critical appraisal is the process of assessing and interpreting evidence by systematically considering its validity, results and relevance to an individual's work. Within the last decade critical appraisal has been added as a topic to many medical school and UK Royal College curricula, and several continuing professional development ventures have been funded to provide further training. This is an update of a Cochrane review first published in 2001.

Objectives

To assess the effects of teaching critical appraisal skills to health professionals on the process of care, patient outcomes and knowledge of health professionals.

Search methods

We updated the search (see [Appendix 1](#) for search strategies by database) and used those search strategies to search the Cochrane Central Register of Controlled Trials (1997 to June 2011) and MEDLINE (from 1997 to June 2011). We also searched EMBASE, CINAHL and PsycINFO (up to January 2010). We searched LISA (up to January 2010), ERIC (up to January 2010), SIGLE (up to January 2010) and Web of Knowledge (up to January 2010). We also searched the Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE) and the Cochrane Effective Practice and Organisation of Care (EPOC) Group Specialised Register up to January 2010.

Selection criteria

Randomised trials, controlled clinical trials, controlled before and after studies and interrupted time series analyses that examined the effectiveness of educational interventions teaching critical appraisal to health professionals. The outcomes included process of care, patient mortality, morbidity, quality of life and satisfaction. We included studies reporting on health professional knowledge/awareness only when based upon objective, standardised, validated instruments. We did not consider studies involving students.

Data collection and analysis

Two review authors independently extracted data and assessed risk of bias. We contacted authors of included studies to obtain missing data.

Main results

In total, we reviewed a total of 11,057 titles and abstracts, of which 148 appeared potentially relevant to the review. We included three studies involving 272 people in this review. None of the included studies evaluated process of care or patient outcomes. Statistically significant improvements in participants' knowledge were reported in domains of critical appraisal (variable approaches across studies) in two of the three studies. We determined risk of bias to be 'unclear' and as such considered this to be 'plausible bias that raises some doubt about the results'.

Authors' conclusions

Low-intensity critical appraisal teaching interventions in healthcare populations may result in modest gains. Improvements to research examining the effectiveness of interventions in healthcare populations are required; specifically rigorous randomised trials employing interventions using appropriate adult learning theories.

PLAIN LANGUAGE SUMMARY**Teaching critical appraisal skills in healthcare settings**

Critical appraisal involves interpreting information, in particular information within research papers, in a systematic and objective manner. This review looked at whether teaching critical appraisal skills to health professionals led to changes in processes of care, patient outcomes, health professionals' knowledge of how to critically appraise research papers, or all three. The review found that teaching critical appraisal skills to health professionals may improve their knowledge. However, there was a lack of good quality evidence as to whether teaching critical appraisal skills led to changes in the process of care or to changes in patient outcomes.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. Teaching critical appraisal for improving process of care variables, patients outcomes and health professionals' knowledge

Teaching critical appraisal for improving process of care variables, patients outcomes and health professionals' knowledge

Patient or population: health professionals

Intervention: teaching critical appraisal

Comparison: usual care

Outcomes	Effects	No of participants (studies)	Quality of the evidence (GRADE)
Process of care	Not reported	-	-
Patient-related outcomes	Not reported	-	-
Knowledge Follow-up: 0 to 6 months	Improvement in knowledge was reported in both studies and was statistically significant	146 (2 studies)	⊕⊕⊕⊕ low ¹
Critical appraisal skills Heterogeneous measurement scales Follow-up: 0 to 6 months	Improvement in critical appraisal skills was reported in all 3 studies. Statistically significant results were found in 2.	160 (3 studies)	⊕⊕⊕⊕ very low ^{1,2}

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

¹ All studies were assessed as unclear risk of bias (unclear risk for one or more key domains) both across and within study assessments, and very few participants in studies.

² Heterogeneity in scales and in significance of effects.

BACKGROUND

There is evidence that findings from research are often either translated slowly into practice or not at all. Such new developments in health care are often communicated through reports of scientific studies being presented at conferences and in academic journals. When healthcare professionals face a dilemma concerning the effectiveness of an intervention for a particular patient, one option is to locate relevant scientific studies. However, reliable interpretation of these retrieved articles requires a basic understanding of scientific and statistical methods, together with the adoption of an inquisitive and sceptical approach. Formal training in these critical appraisal skills may assist healthcare workers in interpreting studies, informing them of potential biases, increasing comprehension of numerical results, and helping them to decide whether articles are relevant, valid and how they should influence the care of their patients. By giving greater understanding of research designs with greater degrees of internal validity, critical appraisal may also assist healthcare workers deal with the recognised and increasing problems of information overload. This not only applies to healthcare professionals but to anyone who is making a decision regarding health care, for example health service managers, users of health care and the media who disseminate information about healthcare issues.

Over the past few decades, there has been a revolution in the teaching of critical appraisal as a tool in medical education across the continuum, in determining the purchase of health care, and in the assessment of higher professional training. There is a need to determine how effective this teaching is, what influences the efficacy of teaching and whether acquisition of critical appraisal skills has any effect on healthcare worker behaviour or patient-related outcomes, bridging the gap between research and practice.

It is clear from general surveillance of the literature that critical appraisal has been subject to development, starting from the point where Sir William Osler and others used journal clubs, "for the purpose and distribution of periodicals to which he could ill afford to subscribe as an individual" in the 19th century (Linzer 1987b). Key stages, the earliest of which would not constitute critical appraisal as we have defined it, include the following.

- Identification of the need to facilitate access to and consideration of research.
- Recognition of the variable quality of research and the need to be able to identify shortcomings, leading in turn to critical appraisal being regarded as the science of trashing research (Fisher 1973; Seltzer 1972).
- Development of checklists to make the process of appraisal more structured, explicit and straightforward (DCEB 1981).
- Emphasis on application of research findings. In particular, changing the main aim from identifying as many things as possible wrong with a piece of research to making a balanced assessment of the strengths as well as the weaknesses of the research and applying it within these limits, in the context of a particular patient or group (EBMWG 1992).
- Firmly embedding critical appraisal in a wider package of skills, e.g. question identification and literature retrieval, that together help the use of research, rather than the expectation that critical appraisal alone will achieve this (Sackett 1997).
- Recognition that the approach is of value to, and can be adapted for use by, all those potentially involved in making healthcare

decisions including: nurses; professions allied to medicine; social workers; healthcare managers; and patients (Milne 1995; Milne 1996).

The important role played in these developments by McMaster University in Canada should be highlighted.

This account of the development of critical appraisal is not complete by any means, but should be sufficient to re-emphasise that critical appraisal is not a static entity. Indeed development continues in many areas, such as extending the approach to a broad range of evidence, e.g. qualitative research and exploring means to improve the structured integration of different disciplines of information into a final decision. A consequence of the continuing evolution of critical appraisal is that precision is required in defining the term 'critical appraisal' in particular. At any point in time it is clear that the term may be being used in different ways by different authors and commentators. Although precise wording varies, there is broad agreement on the definition of critical appraisal being: "The process of assessing and interpreting evidence (usually published research) by systematically considering its validity (closeness to the truth), results and relevance to the individual's work" (Last 1988). This is the definition that we have used in this systematic review.

OBJECTIVES

The review will determine whether teaching critical appraisal improves professional practice or patient outcomes. We will address the following hypotheses:

1. Teaching critical appraisal has an impact on the process of care delivered by healthcare professionals as compared to those who do not receive that teaching.
2. Teaching critical appraisal to healthcare professionals improves patient outcomes resulting from their clinical management compared to patients receiving clinical management from healthcare professionals who have not received critical appraisal teaching.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs), controlled clinical trials (CCTs), controlled before and after studies (CBAs) and interrupted time series (ITSs). The minimal requirement is that there has to be a comparison with no teaching in critical appraisal, either in a separate group or in the same group before teaching was undertaken.

Types of participants

Any qualified healthcare professionals (including managers and purchasers) with direct patient care in any given clinical setting. We excluded studies involving students.

Types of interventions

Educational interventions (defined as a co-ordinated educational activity, of any medium, duration or format) teaching critical appraisal (defined as the process of assessing and interpreting evidence by systematically considering its validity, results and

relevance to ones' own work). The critical appraisal educational intervention could either be a single intervention or one of a package of interventions, as long as data were separately extractable. We judged studies which simply taught biostatistics, or epidemiology, or both as not fulfilling these criteria.

Types of outcome measures

1. Objectively measured process of care variables.
2. Objectively measured patient outcomes (e.g. health outcomes: mortality, morbidity, quality of life and satisfaction).
3. Objectively measured assessments of the impact of teaching critical appraisal on health professionals' knowledge/awareness were considered if assessment of outcome measure was based upon standardised and reliable instruments (tests, questionnaires etc).

Search methods for identification of studies

First version of the review

For the first version of this review, we used electronic and non-electronic searching. We used search strategies to focus searches on design types of particular relevance. We developed subject specific search strategies using text words and index terms gathered in the first instance from existing collections of articles known to be relevant. We reconsidered these search strategies on the basis of the larger pool of relevant articles yielded at the end of the search of the first-line sources. Where the search strategies needed modification, we re-searched bibliographic databases in the first-line sources.

The authors searched the following.

- *The Cochrane Library* to Issue 2 2000, including the Cochrane Controlled Trials Register using: critical appraisal(ft), critical reading(ft), journal club(ft), education-graduate(mh), teach*(mh), continuing medical education(mh), problem based learning(mh), medical education(mh), epidemiol*(mh), reading(mh).
- MEDLINE 1966 to 1997 using: teach*(mh), read*(mh), education*(mh explode), learn*(mh), critical* apprais*(ft)critical* read*(ft), journal* club*(ft).
- EMBASE 1980 to 1997 using: education(mh), teach*(mh), critical* apprais*(ft), critical* read*(ft), learn*(mh).
- ERIC 1966 to 1997 using: critical appraisal(ft), medical education(ft).
- Social Science Citation Index 1981 to 1997 using: critical appraisal (ft), journal club(ft), critical reading(ft).
- Science Citation Index 1981 to 1997 using: critical appraisal (ft), journal club(ft), critical reading(ft). We conducted forward citation searches using the following names taken from leading publications in the field: Linzer, Sackett, Bennett, Riegelman.
- CINAHL 1982 to 1997 using: critical apprais*(ft), critical read*.
- SIGLE 1980 to 1997 using: critical appraisal, evidence based health.
- LISA 1976 to 1997 using terms as in EMBASE.
- PsycLit 1974 to 1997 using terms as in MEDLINE.
- Internet: Webcrawler search engine using search terms (critical appraisal + teaching + medicine).

We made contact with experts in the field directly and through electronic mailing lists. We identified individuals and centres with expertise in teaching critical appraisal through published literature, abstracts of meetings, course advertisements, personal knowledge, local contacts and through networking at relevant meetings (the evidence based educators meetings funded by the UK Department of Health at The King's Fund Centre, London, 1996/1997 and the 5th Annual Cochrane Colloquium, Amsterdam, 1997). We contacted individuals and centres requesting details of published and unpublished studies as well as ongoing work. We requested details of completed and ongoing studies through email discussion lists. We searched medical matrix and pages of institutions and sites known to have interest in critical appraisal, including: School of Health and Related Research (SCHARR), Critical Appraisal Skills Programme (CASP), McMaster University (Hamilton, Ontario, Canada), Centre for Evidence-based Medicine (CEBM) Oxford, and Harvard University websites.

Update of the review

For the update of this review, an information specialist (NS) developed new search strategies from the keywords used in the first version of the review (see [Appendix 1](#)). We performed searches up to January 2010, and then up to June 2011 (in select databases). Specifically, we searched:

- the Ovid Cochrane Central Register of Controlled Trials (CENTRAL) (1997 to 2nd Quarter June 2011);
- Ovid MEDLINE (from 1997 to June 2011);
- Ovid EMBASE, Ovid and Ebscohost CINAHL, and Ovid PsycINFO (from 1997 to January 2010);
- Proquest LISA (from 1976 to January 2010);
- Proquest ERIC (from 1966 to January 2010);
- the Ovid Cochrane Database of Systematic Reviews (CDSR) and Database of Abstracts of Reviews of Effects (DARE), and the Cochrane Effective Practice and Organisation of Care (EPOC) Group Specialised Register (to January 2010); and
- ISI Web of Knowledge (from 1981 to January 2010).

We also reviewed references of relevant articles retrieved from the above search.

Data collection and analysis

Selection of studies

One of two review authors (TH, NS) screened the results of all searches and excluded articles clearly of no relevance to the study. Another author (not the original review author) independently reviewed excluded studies to ensure studies were not excluded inadvertently or inaccurately; this process ensured a conservative approach and that any excluded study was determined in duplicate. Two of the four independent assessors (TH, NS, CH, RS) retrieved and reviewed copies of all other articles, and the assessors judged whether the articles were eligible for inclusion in the review according to the criteria stated above. Negotiation settled any disagreement. We only included abstracts if a full article (published or unpublished) could not be obtained and if the abstract contained enough information to discern the intervention of interest and relevant outcome data.

Assessment of risk of bias in included studies

We used the Cochrane Collaboration's 'Risk of bias' assessment tool (Higgins 2008) that examines the following six criteria: sequence generation, allocation concealment, blinding of participants, personnel and outcome assessors, incomplete outcome data, selective outcome reporting and other sources of bias. Two authors independently assessed criteria and summarised descriptively. We considered conclusions in light of the overall assessments of risk of bias. We planned to use the Cochrane Effective Practice and Organisation of Care (EPOC) Group risk of bias criteria for any eligible non-randomised studies.

We had planned to examine publication bias using a funnel plot but did not find sufficient numbers of studies to run this analysis (N < 10 studies).

Dealing with missing data

We extracted data using a checklist developed using principles outlined by the EPOC Group, amended for the purposes of this review. We reported quantitative data from each study in natural units in a results table. We made contact with lead authors of included trials to request further relevant data not included in the published reports, and to obtain full reports of evaluations that had only been published in abstract form.

Data synthesis

We synthesised all data descriptively using qualitative methods, delineating the validity and size of the study, the exact nature of the intervention (for example length, number of educational sessions, content) and participants in the programme (for example number, profession, experience), and any other differences which might impact on the results.

Quantitative synthesis can only be applied to systematic reviews where the interventions, participants, outcomes and study designs are similar enough to suggest that results can be pooled.

RESULTS

Description of studies

Results of the search

First version of the review

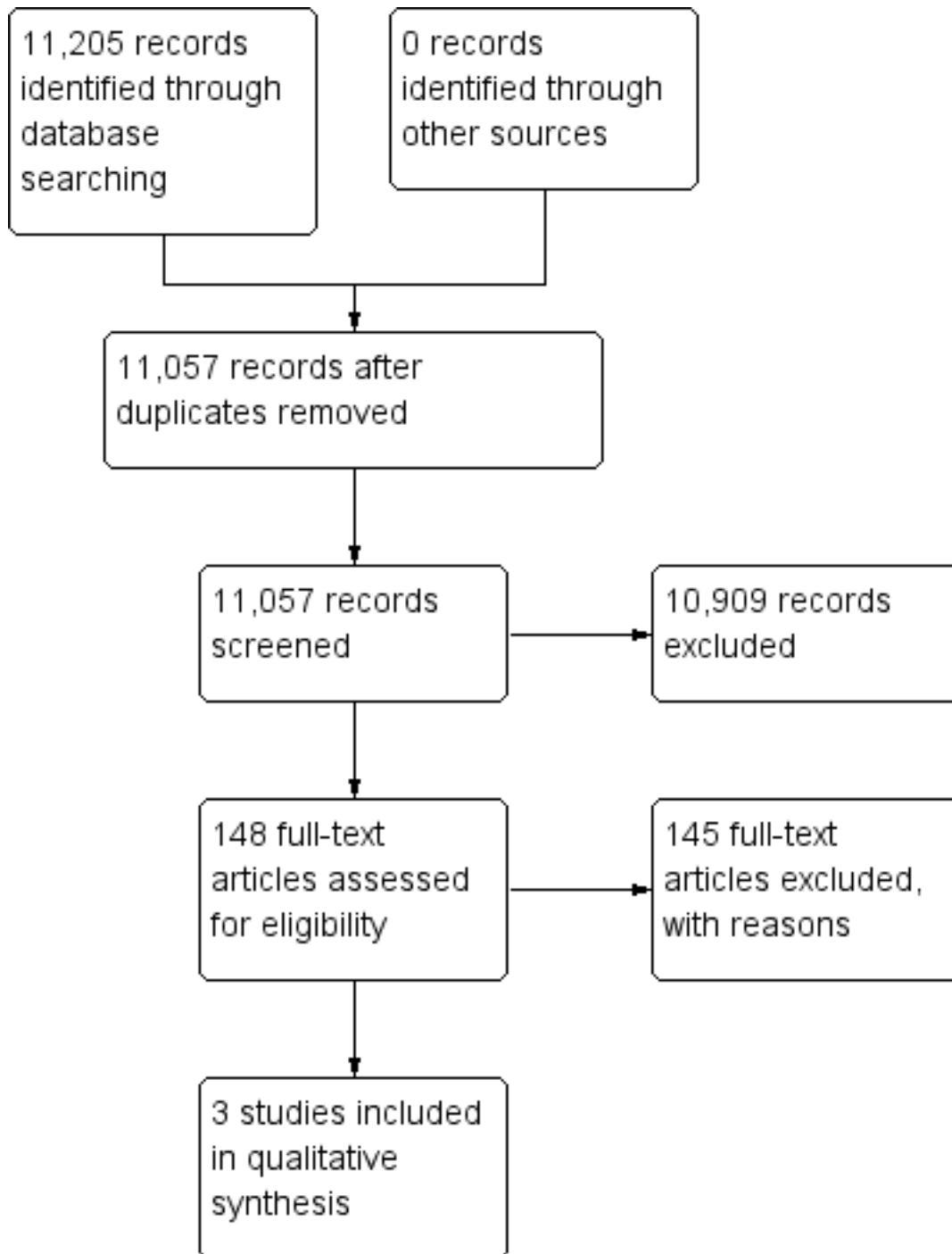
The searches located a total of 4004 articles, of which 137 appeared potentially relevant to the review and were retrieved. We immediately excluded many articles, such as review articles, letters, editorials and articles solely concerned with the process of teaching critical appraisal or of the methodology involved with critical appraisal of the literature. We also excluded descriptive accounts of studies of teaching critical appraisal if they did not include formal evaluations of their impact.

Two independent review authors (RM, JP) formally assessed 47 articles for inclusion in the review. The level of initial agreement on inclusion between the review authors was good, indicating that the inclusion criteria could be reliably applied. All disagreements were resolved by discussion without resorting to a third opinion. We totally excluded 46 of the articles, or excluded them pending further information. We reported full reasons for exclusion in the table [Characteristics of excluded studies](#).

Update of the review

The searches located a total of 11,057 titles and abstracts, of which 148 appeared potentially relevant for the review and full texts were retrieved ([Figure 1](#)).

Figure 1. Study flow diagram.



We included two studies from this updated search (excluding 146 for various reasons including non-qualified healthcare worker population, intervention not including the teaching of critical appraisal, not appropriate etc.). We identified one study as a comparison of two different methods of teaching critical appraisal (McLeod 2010); we will assess this approach in the next update. We reported reasons for exclusion of important studies in the [Characteristics of excluded studies](#) table.

Included studies

First version of the review

One randomised controlled trial met the criteria for inclusion in the review (Linzer 1988).

Update of the review

We included two additional studies, both randomised trials (MacRae 2004; Taylor 2004).

A total of three studies are included in the review.

Summary of included studies

See summarised information within the [Characteristics of included studies](#) table.

One study examined effects in interns ([Linzer 1988](#)) and two in practising physicians ([MacRae 2004](#); [Taylor 2004](#)). Studies were published between 1988 ([Linzer 1988](#)) and 2004 ([MacRae 2004](#); [Taylor 2004](#)). Countries of publication included the United States of America ([Linzer 1988](#)), United Kingdom ([Taylor 2004](#)) and Canada ([MacRae 2004](#)). All three included studies were randomised trial designs. Interventions included a journal club supported by a half-day workshop ([Linzer 1988](#)), critical appraisal materials (package) including papers with methodological reviews, listserve discussions and articles ([MacRae 2004](#)) and a half-day workshop based on a Critical Appraisal Skills Programme (CASP) ([Taylor 2004](#)). Sample size (analysed) ranged between 44 ([Linzer 1988](#)) and 145 ([Taylor 2004](#)) subjects. One study reported that the educational intervention was developed and executed using an adult learning theory ([MacRae 2004](#)).

[Linzer 1988](#) examined whether a journal club improved internal medicine interns reading habits, knowledge of epidemiology and biostatistics, and critical appraisal skills. The study randomised medical interns (N = 44) to receive either journal club or a control seminar series focused on ambulatory medicine; participants in the control group were also assigned to a waiting list to attend the workshop. The general medicine faculty delivered the intervention which included a half-day workshop based on the Critical Appraisal Skills Programme (CASP) (not fully described within the report). Outcomes measured included change in reading habits, knowledge and critical appraisal scores (evaluation instrument assessed baseline characteristics, reading habits, knowledge of epidemiology and biostatistics, and critical appraisal skills). Follow-up evaluation occurred immediately following the intervention (with a mean time between pre-test and post-test scores of 9.5 weeks). The study did not report that the educational intervention was based on any theory-based principles.

[MacRae 2004](#) evaluated the effectiveness of an Internet-based, critical appraisal skills intervention to determine its impact on physician critical appraisal skills. A total of 83 practicing surgeons with no postgraduate training in clinical epidemiology were randomised to a curriculum in critical appraisal skills that included a clinical and methodological article, a listserve discussion, and clinical and methodological critiques or those receiving only the articles. Intervention group participants received eight packages (once monthly containing articles) and questions designed to guide critical appraisal and other supportive resources (listed previously) with the control receiving the eight packages and with online access only to major medical and surgical journals, some of which included articles on critical appraisal. Critical appraisal outcomes were measured using a (locally developed) critical appraisal test. Follow-up evaluation occurred immediately following the intervention. The authors reported that the intervention was based on adult learning theory and findings from continuing education literature.

[Taylor 2004](#) evaluated the effectiveness and costs of a critical appraisal skills educational intervention that was specifically aimed at healthcare professionals. A total of 145 self selected

general practitioners, hospital physicians, professions allied to medicine, and healthcare managers and administrators were randomised to either receive a half-day critical appraisal skills training workshop or a waiting list control. The intervention group received a half-day workshop based on the Critical Appraisal Skills Programme (CASP). Outcomes measured included knowledge of principles for appraising evidence, attitudes towards the use of evidence about health care, evidence-seeking behaviours, and perceived confidence in appraising evidence; critical appraisal skills were assessed through the appraisal of a systematic review. Follow-up evaluation occurred at six months post-intervention. The study did not report that the educational intervention was based on any theory-based principles but did note that approach was based on the model of problem-based small group learning.

Excluded studies

A summary of excluded studies is provided in [Table 1](#).

Risk of bias in included studies

The risk of bias of the included studies is summarised in the 'Risk of bias' tables ([Characteristics of included studies](#)).

Of the included studies, all had 'unclear' risk of bias for one or more key domains ([Linzer 1988](#); [MacRae 2004](#); [Taylor 2004](#)).

Allocation

All studies received 'unclear' for allocation concealment. None of the studies described the method used to conceal the allocation sequence.

Adequate sequence generation

Two of the three studies reported adequately for sequence generation ([MacRae 2004](#); [Taylor 2004](#)). We were unable to determine the method for generating the allocation sequence in one study.

Blinding

Two studies described measures used to blind study participants and personnel from knowledge of which intervention was received ([MacRae 2004](#); [Taylor 2004](#)). The remaining study was described as 'unclear'.

Incomplete outcome data

Only one study sufficiently addressed incomplete outcome reporting ([Linzer 1988](#)). The remaining two studies did not describe the completeness of outcome data for each main outcome and failed to report reasons for attrition/exclusions ([MacRae 2004](#); [Taylor 2004](#)).

Selective reporting

To interpret selective outcome reporting for studies we examined the objective, the intervention and the primary outcomes (one would expect from the study). All studies were deemed to have reported the 'expected outcomes' and those listed as primary outcomes within the methods section. Secondary outcomes were not assessed.

Effects of interventions

See: [Summary of findings for the main comparison Teaching critical appraisal for improving process of care variables, patients outcomes and health professionals' knowledge](#)

Effects data for each study are summarised in the 'Results of included trials' table (Table 2).

1. Does teaching critical appraisal have an impact on the process of care of patients?

Updated version of review: none of the included studies reported process of care outcomes.

2. Does teaching critical appraisal have an impact on patient outcomes?

Updated version of review: none of the included studies reported patient outcomes.

3. Does teaching critical appraisal have an impact on critical appraisal knowledge/awareness of health professionals?

Knowledge scores

Two studies reported knowledge-related outcomes (Linzer 1988; Taylor 2004). Linzer assessed knowledge using an evaluation instrument (given before and immediately following the intervention; interval to follow-up averaged 9.5 weeks) that evaluated baseline research experience, previous education with critical appraisal, reading habits, direct knowledge of epidemiology and biostatistics, and specifically critical appraisal skills (Linzer 1988). The investigators reported that the percentage improvements in knowledge in the two groups are 10% in intervention compared with 2% in control (no P value or confidence interval reported). Additional percentage improvement scores were calculated to adjust for the differences in pre-test scores between the groups, where the numerator is the pre-test/post-test difference and the denominator is the maximum obtainable value minus the pre-test score. Using these adjustments, the intervention group percentage improvement is 26% whilst the control group improvement is 6% ($P = 0.02$). A trend was found that suggested that the more journal club sessions a participant attended, the more knowledge was acquired, resulting in a 'dose-response' relationship in the intervention group that was absent from the control group (intervention $r = 0.40$, $P < 0.07$; control group $r = 0.01$, $P = 0.96$). Taylor 2004 used an 18-question multiple-choice outcome questionnaire focused on attitude and confidence statements. Critical appraisal skills were assessed by the appraisal of a systematic review article independently assessed by two authors. They reported overall knowledge scores as intention-to-treat analysis mean difference (adjusted for sex, age, attendance at previous educational activity, access to medical library, prior experience of searching literature, formal education in research methods and/or epidemiology and/or statistics, and prior involvement in research) as 2.6 (95% confidence interval (CI) 0.6 to 4.6). This was statistically significant at $P \leq 0.05$. An explanatory analysis mean difference was 3.1 (95% CI 1.1 to 5.2), which corresponded to a moderate effect size.

Critical appraisal

All included studies reported critical appraisal-related outcomes. There were significant improvements in domains of critical appraisal (variable approaches across studies) in all but one study

(Taylor 2004). The studies reported outcomes heterogeneously as either mean total scores (Linzer 1988), % improvement (mean) (Linzer 1988), mean total score (as a %) (MacRae 2004), or an estimate of the mean difference (Taylor 2004). Linzer 1988 demonstrated an improvement of 1.5 correct test questions in the intervention group compared to a 0.3 improvement in the control group (mean scores). This translated into a 1.2 correct question difference between intervention and control group which was calculated to be statistically significant between groups ($P = 0.04$). The differences in post-test scores between control and intervention groups was 0.5 correct test questions (3% difference). MacRae 2004 reported the overall mean score from an exam (a locally developed test of critical appraisal) that demonstrated statistically significant differences (mean% (SD)) between the intervention group (58.8% (8)) and the control participants (50% (8)) ($P < 0.01$). Findings from Taylor 2004 resulted in no statistically significant differences observed between the intervention and comparator groups (intention-to-treat mean difference 1.2 (95% CI .01 to 2.4). There were also no differences observed in the ability to appraise methodology or relevance/generalisability of evidence. The authors note that the cost associated with one-off workshops (estimated to be GBP 250) is challenged by their findings.

DISCUSSION

The trials included in this updated review indicate that critical appraisal teaching continues to demonstrate that there is a general improvement in critical appraisal knowledge. The extent and impact of that improvement remains dependent on several factors including whether or not the participants in the intervention groups had some background in research. Most had some prior teaching in critical appraisal therefore resulting in quite high baseline scores (62% of test questions answered correctly, leading to statistical access to 50% of the medical literature). This increased to 60% after the intervention. It may be that the impact of the intervention may have been greater if the baseline performance of the participants had not been as strong. The size of the effect on knowledge attributable to critical appraisal teaching remains difficult to discern. A further challenge beyond estimating the true size of the effect on knowledge is a definition of whether a particular level represents a change, which is likely to be significant in practice.

As has been noted, there is a lack of evidence on the impact of teaching critical appraisal on patient health. However, using patient health measures may not be feasible unless the scale of the critical appraisal programme is very large and sustained with follow-up over many years. It may be more appropriate to concentrate on immediate measures and simultaneously to seek to establish the relationship between such proxy measures and health outcomes through observational data (Mant 1995).

Our review does not examine the impact of teaching critical appraisal skills in undergraduate populations. As stated by Norman 1998, there is empirical evidence to suggest that outcomes of teaching evidence-based medicine (EBM) markedly differ between the undergraduate and postgraduate populations. Norman 1998 suggests that the differences were not explained by methodological shortcomings (e.g. small samples), rather it may be that the integration of EBM as an essential and continuing component of a residency programme will demonstrate much larger and sustained effects that affect patient care. Moreover, from an educationalist

perspective, approaches to learning vary greatly and as such should be considered separately.

Other messages from previous reviews (Audet 1993; Burls 1997; Hyde 2000; Norman 1998; Taylor 2000) include the need for assessment of whether the benefits achievable through investment in time, effort and money required to deliver critical appraisal teaching on a large scale are of broadly similar size to what may be achieved through an equivalent amount spent directly on patient care. More recently, a review by Coomarasamy and colleagues evaluated the effects of standalone versus clinically integrated teaching in evidence-based medicine and found that standalone teaching improved knowledge but not skills, attitudes or behaviours (Coomarasamy 2004). The traditional view of professional learning (that which sees healthcare professionals attending learning workshops/courses/conferences) undervalues the importance of what we learn from our daily practice routines and perpetuates the recommendations that 'teaching' of EBM should be done within a classroom and not in clinical practice.

Although most reviews on EBM have an overall positive effect of teaching critical appraisal on knowledge gains and skills, it should be mentioned that most have been broader in scope and have included all healthcare workers, and examined knowledge, skills, behaviour and attitudes outcomes. Previous reviews have suggested that randomised controlled trials (RCTs) are the research design of choice for assessing the effects and effectiveness of teaching critical appraisal to avoid the threats to validity of past research. This is generally supported by this review although restriction solely to RCTs may be too dogmatic. There are circumstances where valid outcome measures may be collected from before and after studies. These may be where the outcome is objective, the intervention short and outcome assessed immediately after the intervention (with exact timing of measurement being clearly reported); any change may be reasonably attributed to the intervention. Loss to follow-up may not be more difficult in before and after designs, and again this would minimise bias.

To estimate the effectiveness of interventions for improving health professionals' critical appraisal skills, at a minimum it will be important to conduct well-designed, methodologically rigorous, clearly described studies. If interventions are to be supported and taken up by academic health centres, hospitals, medical schools etc. it will be imperative to demonstrate the costs (both monetary and time) of the interventions; none of the studies included in our review examined costs. To improve our understanding of the interventions and to build on previous research it will be important for investigators and study authors to consider adult learning theories and models when developing educational interventions; no included studies reported that educational interventions were designed or supported by a particular theory. To ensure reliability and validity, future investigators should use valid and reliable outcome measures. Independent, duplicate, direct observation is recommended by individuals trained to assess the given outcomes.

The objectives of this review were limited to the evaluation of the impact of critical appraisal teaching on process of care variables, patient outcomes and critical appraisal knowledge of qualified healthcare workers. Therefore, it is not an evaluation of critical appraisal skills programmes, centres for evidence-based health care, clinical effectiveness, clinical governance or any of the more 'macro' uses to which this 'micro' intervention might be harnessed

(Wilkes 1999). This is a fast-moving field (Sackett 1998), therefore evaluations of teaching programmes may be turning to the issue of how, rather than whether, to teach critical appraisal skills.

Educational research has not yet acquired the prestige of pure or applied medical science and may attract inadequate funding, with little effect on career advancement, thus inspiring temporary enthusiasms (Hutchinson 1999). All of these may help to explain the paucity of the literature. This picture will have to change if any of the outstanding questions are to be answered by high-quality research.

Finally, absence of evidence is not evidence of absence. This systematic review has highlighted the poor evidence in the area but it is not evidence for discontinuing critical appraisal teaching.

AUTHORS' CONCLUSIONS

Implications for practice

Results from two included trials suggest that critical appraisal teaching may have positive effects on participants' knowledge. None of the studies evaluated the effect of critical appraisal on the process of health care or on patient health. It is unclear whether the size of benefit in knowledge is large enough to be of practical significance. This review provides some reassurance to those who have invested in critical appraisal teaching activities in that they are likely to have a positive impact. Due to limitations on validity and significance in practice and the absence of results for important outcomes, the evidence is not sufficient to encourage further expansion of critical appraisal activities without inclusion of rigorous evaluations of effectiveness that include, but are not limited to, randomisation.

Implications for research

On the basis of this review the top priority should be a high-quality, multicentre randomised controlled trial of teaching critical appraisal to postgraduates as part of continuing professional development. It should define in advance changes in outcomes that are 'significant in practice'. It should include a consideration of costs and attempt to answer the value for money question. There are many centres around the world teaching critical appraisal to their healthcare students and postgraduate workers. It must be possible to plan and even co-ordinate the work done in this field to increase study numbers and obtain valid results of measurable outcomes, using consistent, valid instruments. This could facilitate the production of more homogeneous results, which could be combined to give an overall effect direction and size of effect. One way of achieving this might be multicentre, methodologically rigorous, controlled before and after studies in comparable groups with the same valid instruments, measuring the same outcomes, which may be able to give answers of reasonable validity when randomisation is not feasible. It would also be important to note that only one study acknowledged the application of adult learning theory principles when designing its curriculum. It is imperative that, moving forward, interventions are designed and developed considering these principles (e.g. motivations, barriers) that are particular to the learner.

ACKNOWLEDGEMENTS

We gratefully acknowledge the contributions made by Ms. Danielle Rabb, MLIS towards reviewing risk of bias assessments, data extractions, and reviewing critically and editorially the final tables

and manuscript. We also would like to thank Prof Jonathan J Deeks for his contribution to an earlier version of this review.

REFERENCES

References to studies included in this review
Linzer 1988 {published data only}

Linzer M, Brown JT, Frazier LM, DeLong ER, Siegel WC. Impact of a medical journal club on house-staff reading habits, knowledge, and critical appraisal skills. A randomized control trial. *JAMA* 1988;**260**(17):2537-41. [MEDLINE: 89012401]

MacRae 2004 {published data only}

MacRae HM, Regehr G, McKenzie M, Henteleff H, Taylor M, Barkun J, et al. Teaching practicing surgeons critical appraisal skills with an Internet-based journal club: a randomized, controlled trial. *Surgery* 2004;**136**:641-6.

Taylor 2004 {published data only}

Taylor RS, Reeves BC, Ewings PE, Taylor RJ. Critical appraisal skills training for health care professionals: a randomized controlled trial. *BMC Medical Education* 2004;**4**(30):1-10.

References to studies excluded from this review
Alarcón 1994 {published data only}

Alarcón PE, Velázquez LV. Effect of an educational intervention in the critical reading of medical students [Efecto de una estrategia educativa sobre la lectura crítica de estudiantes de medicina]. *Revista de Investigacion Clinica* 1994;**46**:447-56.

Bennett 1987 {published data only}

Bennett KJ, Sackett DL, Haynes RB, Neufeld VR, Tugwell P, Roberts R. A controlled trial of teaching critical appraisal of the clinical literature to medical students. *JAMA* 1987;**257**(18):2451-4. [MEDLINE: 87199102]

Burls 1997 {published data only}

Burls AJE. An evaluation of the impact of half-day workshops teaching critical appraisal skills. Oxford: Institute of Health Sciences, 1997.

Burstein 1996 {published data only}

Burstein JL, Hollander JE, Barlas D. Enhancing the value of journal club: use of a structured review instrument. *American Journal of Emergency Medicine* 1996;**14**(6):561-3. [MEDLINE: 97010722]

Caudill 1993 {unpublished data only}

Caudill TS, Johnson MMS, Rich EC. The effect of journal club on medicine resident interaction with original research literature. *Clinical Research* 1993;**41**(4):816A.

Cuddy 1984 {published data only}

Cuddy PG, Elenbaas JK, Coit KJ. The effectiveness of a slide-tape program on literature evaluation. *Journal of Biological Communications* 1984;**11**(3):2-4. [MEDLINE: 85054769]

Domholdt 1994 {published data only}

Domholdt E, Flaherty JL, Phillips JM. Critical appraisal of research literature by expert and inexperienced physical therapy researchers. *Physical Therapy* 1994;**74**(9):853-60. [MEDLINE: 94344986]

Dorsch 1990 {published data only}

Dorsch JL, Frasca MA, Wilson MA, Tomsic ML. A multidisciplinary approach to information and critical appraisal instruction. *Bulletin of the Medical Library Association* 1990;**78**(1):38-44. [MEDLINE: 90105958]

Fowkes 1984 {published data only}

Fowkes FG, Gehlbach SH, Farrow SC, West RR, Roberts CJ. Epidemiology for medical students: a course relevant to clinical practice. *International Journal of Epidemiology* 1984;**13**(4):538-41. [MEDLINE: 85104025]

Frasca 1992 {published data only}

Frasca MA, Dorsch JL, Aldag JC, Christiansen RG. A multidisciplinary approach to information management and critical appraisal instruction: a controlled study. *Bulletin of the Medical Library Association* 1992;**80**(1):23-8. [MEDLINE: 92163728]

Gehlbach 1980 {published data only}

Gehlbach SH, Bobula JA, Dickinson JC. Teaching residents to read the medical literature. *Journal of Medical Education* 1980;**55**(4):362-5. [MEDLINE: 80162771]

Gehlbach 1985 {published data only}

Gehlbach SH, Farrow SC, Fowkes FG, West RR, Roberts CJ. Epidemiology for medical students: a controlled trial of three teaching methods. *International Journal of Epidemiology* 1985;**14**(1):178-81. [MEDLINE: 85181837]

Globerman 1993 {published data only}

Globerman J. Teaching critical appraisal of the social work literature. *Journal of Teaching in Social Work* 1993;**7**(2):63-80.

Griffith 1988 {published data only}

Griffith OW, Goss ME, Tardiff K, Cassell EJ, Pardee JD, Weinstein AM. An elective course for first-year students based on the New England Journal of Medicine. *Journal of Medical Education* 1988;**63**(7):559-61. [MEDLINE: 88259186]

Harewood 2009 {published data only}

Harewood GC, Hendrick LM. Prospective, controlled assessment of the impact of formal evidence-based medicine teaching workshop on ability to appraise the medical literature. *Irish Journal of Medical Science* 2010;**179**(1):91-4.

Heiligman 1991 {published data only}

Heiligman RM. Resident evaluation of a family practice residency journal club. *Family Medicine* 1991;**23**(2):152-3. [MEDLINE: 91244109]

Heller 1984 {published data only}

Heller RF, Peach H. Evaluation of a new course to teach the principles and clinical applications of epidemiology to medical students. *International Journal of Epidemiology* 1984;**13**(4):533-7.

Herbert 1990 {published data only}

Herbert CP. Teaching prevention by debate. *Family Medicine* 1990;**22**(2):151-3.

Hicks 1994 {published data only}

Hicks C. Bridging the gap between research and practice: an assessment of the value of a study day in the developing critical research reading skills in midwives. *Midwifery* 1994;**10**(1):18-25. [MEDLINE: 94211153]

Hillson 1993 {unpublished data only}

Hillson SD, Schlossberg LA. Effects of journal club on critical review. *Clinical Research* 1993;**41**(2):559A.

Ibbotson 1998 {published data only}

Ibbotson T, Grimshaw J, Grant A. Evaluation of a programme of workshops for promoting the teaching of critical appraisal skills. *Medical Education* 1998;**32**(5):486-91. [MEDLINE: 99227735]

Inui 1981 {published data only}

Inui TS. Critical reading seminars for medical residents. Report of a teaching technique. *Medical Care* 1981;**19**(1):122-4. [MEDLINE: 81121205]

Johnson 1995 {published data only}

Johnson JM, Reineck C, Daigle-Bjerke A, Goupil NM, Captain C. Understanding research articles. A pilot study of critical reading of research publications. *Journal of Nursing Staff Development* 1995;**11**(2):95-9. [MEDLINE: 95239377]

Kerrison 1995 {published data only}

Kerrison S, Rosenthal J, Wallace P. An evaluation of the North Thames Research Appraisal Teaching project. London. Department of General Practice and Primary Care, Royal Free Hospital School of Medicine June 1995.

Kitchens 1989 {published data only}

Kitchens JM, Pfeifer MP. Teaching residents to read the medical literature: a controlled trial of a curriculum in critical appraisal/clinical epidemiology. *Journal of General Internal Medicine* 1989;**4**(5):384-7. [MEDLINE: 90011416]

Konen 1990 {published data only}

Konen JC, Fromm BS. A family practice residency curriculum in critical appraisal of the medical literature. *Family Medicine* 1990;**22**(4):284-7. [MEDLINE: 90346266]

Krueger 2006 {published data only}

Krueger PM. Teaching critical appraisal: a pilot randomized controlled outcomes trial in undergraduate osteopathic medical education. *Journal of the American Osteopathic Association* 2006;**106**(11):658-62.

Landry 1994 {published data only}

Landry FJ, Pangaro L, Kroenke K, Lucey C, Herbers J. A controlled trial of a seminar to improve medical student attitudes toward knowledge about, and use of the medical literature. *Journal of General Internal Medicine* 1994;**9**(8):436-9. [MEDLINE: 95054564]

Langkamp 1992 {published data only}

Langkamp DL, Pascoe JM, Nelson DB. The effect of a medical journal club on residents' knowledge of clinical epidemiology and biostatistics. *Family Medicine* 1992;**24**(7):528-30. [MEDLINE: 93012638]

Linzer 1987a {published data only}

Linzer M, DeLong ER, Hupart KH. A comparison of two formats for teaching critical reading skills in a medical journal club. *Journal of Medical Education* 1987;**62**(8):690-2. [MEDLINE: 87283776]

MacAuley 1996 {published data only}

MacAuley D. Critical reading using the READER acronym at an international workshop. *Family Practice* 1996;**13**(1):104-5. [MEDLINE: 96252742]

MacAuley 1997 {published data only}

* MacAuley D, Sweeney KG. Critical reading using the READER acronym by experienced general practitioners (GPs) and by GP registrars in southern and Northern Ireland. *Irish Journal of Medical Science* 1997;**166**(3):121-3. [MEDLINE: 97400657]

MacAuley 1998 {published data only}

MacAuley D, McCrum E, Brown C. Randomised controlled trial of the READER method of critical appraisal in general practice. *BMJ* 1998;**316**:1134-9.

MacAuley 1999 {published data only}

MacAuley D, McCrum E. Critical appraisal using the READER method: a workshop-based controlled trial. *Family Practice* 1999;**16**:90-3.

Markert 1989 {published data only}

Markert RJ. A research methods and statistics journal club for residents. *Academic Medicine* 1989;**64**(4):223-4. [MEDLINE: 89166017]

Milne 1996 {published data only}

Milne R, Oliver S. Evidence-based consumer health information: developing teaching in critical appraisal skills. *International Journal of Quality in Health Care* 1996;**8**(5):439-45. [MEDLINE: 97135625]

Mulvihill 1973 {published data only}

Mulvihill MN, Banta HD. A clinical approach to teaching epidemiology and biostatistics. *Journal of Medical Education* 1973;**48**(6):591-3. [MEDLINE: 73178192]

Mulvihill 1980 {published data only}

Mulvihill MN, Wallman G, Blum S. A seven-year retrospective view of a course in epidemiology and biostatistics. *Journal of Medical Education* 1980;**55**(5):457-9. [MEDLINE: 80206902]

Novick 1985 {published data only}

Novick LF, Greene C, Vogt RL. Teaching medical students epidemiology: utilizing a state health department. *Public Health Reports* 1985;**100**(4):401-5. [MEDLINE: 85271101]

O'Sullivan 1995 {published data only}

O'Sullivan P, Pinsky J, Jeremiah J, Wartman S. A learner centred journal club. *Teaching and Learning in Medicine* 1995;**7**(2):121-4.

Radack 1986 {published data only}

Radack KL, Valanis B. Teaching critical appraisal and application of medical literature to clinical problem-solving. *Journal of Medical Education* 1986;**61**(4):329-31. [MEDLINE: 86171692]

Reineck 1995 {published data only}

Reineck CA. Pilot testing an independent study series on research. *Journal of Continuing Education In Nursing* 1995;**26**(6):249-52. [MEDLINE: 96080497]

Riegelman 1986 {published data only}

Riegelman RK. Effects of teaching first-year medical students skills to read medical literature. *Journal of Medical Education* 1986;**61**(6):454-60. [MEDLINE: 86227430]

Romm 1989 {published data only}

Romm FJ, Dignan M, Herman JM. Teaching clinical epidemiology: a controlled trial of two methods. *American Journal of Preventive Medicine* 1989;**5**(1):50-1. [MEDLINE: 89302782]

Salmi 1991 {published data only}

Salmi LR, Collet JP. Critical reading of medical papers. I. Introduction [Lecture critique des articles médicaux]. *Revue du Praticien* 1991;**41**(25):2598-605. [MEDLINE: 92205225]

Sandifer 1996 {published data only}

Sandifer Q.D, Lo SV, Crompton PG. Evaluation of a journal club as a forum to practise critical appraisal skills. *Journal of the Royal College Physicians of London* 1996;**30**(6):520-2. [MEDLINE: 97120534]

Seelig 1991 {published data only}

Seelig CB. Affecting residents' literature reading attitudes, behaviors, and knowledge through a journal club intervention. *Journal of General Internal Medicine* 1991;**6**(4):330-4. [MEDLINE: 91366439]

Seelig 1993 {published data only}

Seelig CB. Changes over time in the knowledge acquisition practices of internists. *Southern Medical Journal* 1993;**86**(7):780-3. [MEDLINE: 93310553]

Stern 1995 {published data only}

Stern DT, Linzer M, O'Sullivan PS, Weld L. Evaluating medical residents' literature appraisal skills. *Academic Medicine* 1995;**70**(2):152-4. [MEDLINE: 95169179]

Viniegra 1994 {published data only}

Viniegra L, Espinosa P. Critical reading in select groups of medical students [Lectura critica en grupos escogidos de estudiantes de medicina]. *Revista de Investigacion Clinica* 1994;**46**(5):407-15. [MEDLINE: 95140946]

References to studies awaiting assessment
McLeod 2010 {unpublished data only}

McLeod RS, MacRae HM, McKenzie ME, Victor JC, Brasel KJ, Evidence Based Reviews in Surgery Steering Committee. A moderated journal club is more effective than an Internet journal club in teaching critical appraisal skills: results of a multicenter randomized controlled trial. *Journal of the American College of Surgeons* 2010;**211**(6):769-76.

Additional references
Audet 1993

Audet N, Gagnon R, Ladouceur R, Marcil M. How effective is the teaching of critical analysis of scientific publications? Review of studies and their methodological quality [L'enseignement de l'analyse critique des publications scientifiques médicales est-il efficace? Révision des études et de leur qualité méthodologique]. *CMAJ* 1993;**148**(6):945-52. [MEDLINE: 93193031]

Coomarasamy 2004

Coomarasamy A, Khan KS. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. *BMJ* 2004;**329**(7473):1017. [PUBMED: 15514348]

DCEB 1981

Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre. How to read clinical journals. I. Why to read them and how to start reading them critically. *CMAJ* 1981;**124**(5):555-8. [MEDLINE: 81135384]

EBMWG 1992

Evidence-based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992;**268**(17):2420-5. [MEDLINE: 93021625]

Fisher 1973

Fisher B. Cooperative clinical trials in primary breast cancer: a critical appraisal. *Cancer* 1973;**31**(5):1271-86. [MEDLINE: 73180231]

Higgins 2008

Higgins JPT, Altman DG (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins, JPT, Green S (editors). *Cochrane Handbook of Systematic Reviews of Interventions Version 5.0.1* [updated September 2008], The Cochrane Collaboration. Available from www.cochrane-handbook.org.

Hutchinson 1999

Hutchinson L. Evaluating and researching the effectiveness of educational interventions. *BMJ* 1999;**318**(7193):1267-9. [MEDLINE: 99247962]

Hyde 2000

Hyde C, Parkes J, Deeks J, Milne R. Systematic review of studies of teaching critical appraisal. UK National R&D Programme: Evaluating methods to promote the implementation of R&D. Project Reference 12-8 2000.

Last 1988

Last JM (ed). *A Dictionary of Epidemiology*. New York: Oxford University Press, 1988.

Linzer 1987b

Linzer M. The journal club and medical education: over one hundred years of unrecorded history. *Postgraduate Medical Education* 1987;**63**(740):475-8. [MEDLINE: 88124546]

Mant 1995

Mant J, Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ* 1995;**311**(7008):793-6. [MEDLINE: 96061751]

Milne 1995

Milne R, Donald A, Chambers L. Piloting short workshops on the critical appraisal of reviews. *Health Trends* 1995-96;**27**(4):120-3. [MEDLINE: 97071624]

Norman 1998

Norman GR, Shannon SI. Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: a critical appraisal. *CMAJ* 1998;**158**(2):177-81. [MEDLINE: 98130390]

Sackett 1997

Sackett DL, Richardson WS, Rosenberg WM, Haynes RB. *Evidence-based medicine. How to practice and teach EBM*. Edinburgh: Churchill Livingstone, 1997.

Sackett 1998

Sackett DL, Parkes J. Teaching critical appraisal: no quick fixes. *CMAJ* 1998;**158**(2):203-4. [MEDLINE: 98130393]

Seltzer 1972

Seltzer CC. Critical appraisal of the Royal College of Physicians' report on smoking and health. *Lancet* 1972;**1**(7744):243-8. [MEDLINE: 72096974]

Taylor 2000

Taylor R, Reeves B, Ewings P, Binns S, Keast J, Mears R. A systematic review of the effectiveness of critical appraisal skills training for clinicians. *Medical Education* 2000;**34**(2):120-5. [MEDLINE: 20118035]

Wilkes 1999

Wilkes M, Bligh J. Evaluating educational interventions. *BMJ* 1999;**318**(7193):1269-72. [MEDLINE: 99247963]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Linzer 1988

Methods	Trial design: RCT Power calculation: Y Funding: Kate B. Reynolds Health Care Trust Ethics approval: not reported Lost to follow-up: 3
Participants	Characteristics of participants: all interns entering the Department of Internal Medicine at Duke University, Durham, NC Exclusion criteria: not reported No. randomised: 44 No. analysed: 41 Other characteristics: age: not reported; sex: not reported Setting: hospital (Department of Internal Medicine at Duke University, Durham, NC) Country: USA
Interventions	Intervention group: medical journal club led by interns using 1 article (provided critique methods, results, conclusions); presentation supported by general medicine faculty facilitator (elaborates on epidemiological/biostatistical concepts). Typically involved 1-hour of preparation with half of that hour on selecting an article.

Linzer 1988 (Continued)

Theory-based intervention: N

Control group: standard conference series on ambulatory medicine

Intervention deliverer: general medicine faculty

Outcomes

Outcomes measured: change in reading habits; knowledge and critical appraisal scores (evaluation instrument assessed: baseline characteristics, reading habits, knowledge of epidemiology and biostatistics, and critical appraisal skills). A non-validated evaluation instrument (given before and immediately following the intervention) that evaluated baseline research experience, previous education with critical appraisal, reading habits, direct knowledge of epidemiology and biostatistics, and specifically critical appraisal skills

Distal follow-up interval: immediately following the educational intervention (average time between pre-test and post-test evaluation, 9.5 weeks)

Losses to follow-up: 3

N randomised: 22 IG, 22 CG

N completing follow-up: 22 IG, 19 CG

Reasons for loss to follow-up: CG - 1 control subject declined the pre-test, one medicine-pediatrics resident was inappropriately entered into the study (control group), and 1 control group intern left the training programme

Notes
Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	44 interns entering the Department of Internal Medicine at Duke University, Durham, NC, were randomly assigned to either a general medicine journal club or a standard conference in topics in ambulatory medicine
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	(Unclear for consent forms only): the specific study hypotheses and, in particular, the focus on impact of the journal club on these matters were not noted in the consent form (Yes for outcome assessors): for the final grading process, graders were blinded to intervention group and to whether it was a pre-test or post-test (p.2538)
Incomplete outcome data (attrition bias) All outcomes	Low risk	One control subject declined the pre-test, one medicine-pediatrics resident was inappropriately entered into the study (control group), and 1 control group intern left the training programme
Selective reporting (reporting bias)	Low risk	Expected outcomes reported
Other bias	Low risk	

MacRae 2004
Methods

Trial design: RCT

Teaching critical appraisal skills in healthcare settings (Review)

Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

MacRae 2004 (Continued)

Power calculation: N

Funding: Physician's Services Incorporated, Ethicon and Ethicon Endosurgery

Ethics approval: Y

Lost to follow-up: 2

Participants

Characteristics of participants: active members of the Canadian Association of General Surgeons who had access to the Internet and to e-mail, and agreed to be randomised, would complete a written exam

Exclusion criteria: surgeons with postgraduate training in clinical epidemiology

No. randomised: 81

No. analysed: 55

Other characteristics: age: not reported; sex: not reported

Setting: Internet-based

Country: Canada

Interventions

Intervention group: participants in the intervention group received 8 packages (of 2 articles from 1 October 2001 to 1 May 2002) and questions designed to guide critical appraisal. 1-week listserv discussions centred on methodology of articles, a comprehensive methodologic review of the article prepared by a surgeon with training in clinical epidemiology and a clinical review were also provided. Articles chosen to be of clinical interest to the general surgeons and the methodological article that directly related to the clinical article. This was an internet-based intervention.

Theory-based intervention: Y (based on adult learning theory and findings from continuing education literature)

Control group: received the 8 packages (of 2 articles), clinical articles and had online access to major medical and surgical journals, some of which included articles on critical appraisal

Intervention deliverer: sessions moderated by a general surgeon with training in clinical epidemiology at the level of a master of science degree

Outcomes

Outcomes measured: (locally developed) critical appraisal used previously on a group of general surgery residents (internal consistency Chronbach alpha 0.77) with evidence of face, content and construct validity. Inter-rater reliability of 2 physicians marking the examination was 0.93. Participants completed rating scales on aspects of study quality and gave free-text responses to questions about methodology, the validity and applicability of results. Two general surgeons developed an answer key and all answers were marked by a single clinical epidemiologist.

Distal follow-up interval: immediately following intervention. Examinations were mailed to participants in both groups with an e-reminder 3 weeks later. Asked to submit within 6 weeks.

Losses to follow-up: 26

N randomised: 44 IG, 37 CG

N completing follow-up: 26 IG, 29 CG

Reasons for loss to follow-up: 3 were excluded prior to randomisation, and another 2 were excluded after randomisation (1 in each group) because they had previous clinical epidemiology training. Fewer of the experimental group participants (26 surgeons, 58%) than the control group (29 surgeons, 76%) completed the examination. This difference may have been due to trial fatigue, in that we asked the experimental group to complete much more material on a monthly basis. Toward the end of the 8 months, the average number of returned evaluations also decreased, again suggesting the possibility of trial fatigue. (p. 644)

MacRae 2004 (Continued)

Notes

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Participants were stratified first by type of practice (academic versus community) then by year of graduation. Within type of practice, participants were blocked into groups of 10 based on year of graduation. This process ensured random assignment within the constraints (1) that exactly half of the academic and exactly half of the non-academic participants who started the study were assigned to each condition and (2) that, within academic and non-academic groups, year of graduation would not be confounded with condition. (p.642)
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	The examinations were all marked by a single clinical epidemiologist. All identifying data were blacked out on the examination prior to marking by the research assistant. The control and intervention group examinations were mixed and then marked in random order by a clinical epidemiologist. (p.643)
Incomplete outcome data (attrition bias) All outcomes	High risk	Not reported
Selective reporting (reporting bias)	Low risk	Expected outcomes reported
Other bias	High risk	The major potential bias that may have affected our result was that participants were recruited on a volunteer basis, and are thus more motivated, and perhaps more likely to benefit from the intervention than the general population of surgeons

Taylor 2004

Methods	<p>Trial design: RCT</p> <p>Power calculation: Y</p> <p>Funding: NHS R&D Executive: evaluating methods to practice the implementation of R&D (project no. IMP 12-9)</p> <p>Ethics approval: Y</p> <p>Lost to follow-up: questionnaire: 40; critique: 81</p>
Participants	<p>Characteristics of participants: healthcare practitioners working within the South and West Regional Health Authority</p> <p>Exclusion criteria: attendance at a previous Critical Appraisal Skills (CAS) workshop</p> <p>No. randomised: 145</p> <p>No. analysed: questionnaire: 105, critique: 64</p> <p>Other characteristics: age: (majority of participants) 40 to 49 years: IG 37 (50.6%), CG 32 (44.4%); sex: IG: M 48, F 25, CG: M 46, F 26</p>

Taylor 2004 (Continued)

Setting: workshop

Country: England

Interventions

Intervention group: half-day workshop based on Critical Appraisal Skills Programme (CASP). The 3-hour workshop is focused on facilitating a process whereby research evidence is systematically examined to assess study validity, the results and relevance to a clinical scenario. The workshop begins with an introductory talk (~60 minutes) about the importance of evidence-based healthcare practice, theoretical basis for systematic reviews and the JAMA appraisal guideline. Small group work (~30 minutes) and then a plenary session run by 3 to 4 individuals each with formal training in health services research methods.

Theory-based intervention: N

Control group: assigned to a waiting list to attend a workshop

Intervention deliverer: not reported

Outcomes

Outcomes measured: (using a validated tool) 18 multiple-choice questions focused on knowledge of principles for appraising evidence, attitudes towards the use of evidence about healthcare, evidence seeking behaviours, and perceived confidence in appraising evidence; critical appraisal skills were assessed through the appraisal of a systematic review.

Distal follow-up interval: 6 months (post intervention)

Losses to follow-up: questionnaire: 40; critique: 81

N randomised: 72 IG, 73 CG

N completing follow-up: questionnaire: 44 IG, 61 CG; critique: 21 IG, 43 CG

Reasons for loss to follow-up: not reported

Notes

It is plausible that respondents may have differed in some way to non-respondents, such as in their level of motivation, and may therefore responded more positively to this educational intervention. However, this was not supported by the poor outcome response rate.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	An independent researcher used computer-generated codes to allocate applicants randomly to intervention or control group (p.2)
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding (performance bias and detection bias) All outcomes	Low risk	An independent researcher used computer-generated codes to allocate applicants randomly to intervention or control group (p.2) The researchers who scored study outcomes were blinded to the allocation of participants at all times. (p.2)
Incomplete outcome data (attrition bias) All outcomes	High risk	Not reported
Selective reporting (reporting bias)	Low risk	Expected outcomes reported

Taylor 2004 (Continued)

Other bias	High risk	The number of participants recruited was less than that intended, not all participants provided outcomes and the trial was about 20 per cent under the desired power.
------------	-----------	---

CG: control group

IG: intervention group

 JAMA: *Journal of the American Medical Association*

RCT: randomised controlled trial

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Alarcón 1994	The study was provisionally excluded pending clarification of the nature of the intervention, which did not clearly seem to consist of teaching critical appraisal. Further, it was not clear what the control group had received. Unfortunately, no clarification was received and the study was effectively excluded, despite clearly being of potential relevance.
Bennett 1987	This controlled trial of 92 medical students compared 8 seminars on critical appraisal with control group measuring skills. This was excluded as it did not teach qualified healthcare workers.
Burls 1997	This was a multidisciplinary 1-day workshop teaching critical appraisal to 1880 healthcare workers, and evaluating self assessed knowledge and attitudes. The study was excluded as it did not assess outcomes using objective measures.
Burstein 1996	This before and after study in emergency medicine residents assessed the impact of a structured review instrument in a journal club on overall satisfaction, perceived educational value, attendance and workload. It was excluded as the instrument did not constitute teaching of critical appraisal.
Caudill 1993	This before and after study evaluated a comprehensive multi component course in critical appraisal skills for medical residents. Knowledge, attitudes and behaviour were the outcomes assessed after the 6-month course, with knowledge outcomes being assessed by objective validated measures (same as in Linzer RCT). This study was excluded as it did not have a control comparison group and did not meet other inclusion criteria for study methodology.
Cuddy 1984	This controlled trial of 18 medical students involved the evaluation of slide tape presentation/lecture teaching critical appraisal. It was excluded as it did not teach qualified healthcare workers, and had not stated the validity of instruments used.
Domholdt 1994	This study compared the critical appraisal skills of inexperienced and expert physical therapists and examined factors influencing critical appraisal skills. It was excluded as the impact of teaching of critical appraisal was not evaluated.
Dorsch 1990	Description of the development of a critical appraisal teaching intervention for the study by Frasca. It was excluded because it was duplicative.
Fowkes 1984	This before and after study evaluated an epidemiology course for medical students which included teaching critical appraisal. It was excluded on confirmation that there was no separable data for that part of the course teaching critical appraisal.
Frasca 1992	This was a controlled trial of 92 medical students comparing 10 seminars on critical appraisal teaching and assessing skills. This was excluded as it did not teach qualified healthcare workers.

Study	Reason for exclusion
Gehlbach 1980	This controlled trial of 35 US family medicine residents evaluated 8 seminars of critical appraisal teaching with examination at end of year. It was excluded as validation of instrument was not stated.
Gehlbach 1985	This controlled trial compared teaching an 'Epidemiology for Clinical Practice' course delivered by lectures, small group seminars or self learning packages. It was excluded because there was no comparison with critical appraisal teaching. It was identified as a potentially useful study outside the immediate context of this review.
Globerman 1993	This describes the development of a course to teach critical appraisal skills to social workers. The main focus is descriptive, but there is reference to evaluation. It was excluded because there was no comparison group and no pre-course assessment.
Griffith 1988	This study reported a course that focused on teaching critical appraisal skills to medical students. It was excluded because there was no comparison group and no pre-course assessment.
Harewood 2009	This study was a 1 group pre-post study
Heiligman 1991	This study identified attitudes of family practice residents toward a journal club and the identification of factors contributing to the success of the journal club. It was excluded because there was no comparison group and no pre-course assessment.
Heller 1984	This controlled trial assessed the impact of a general epidemiology and public health course to medical students. Although the course included the development of a critical approach to information, it was excluded because it was judged to be a small component, the effect of which could not be separated.
Herbert 1990	This study reported a course teaching critical appraisal to residents and its application in an innovative debate format for making clinical decisions. It was excluded because no comparative evaluative data were presented, the absence of which was confirmed by the author.
Hicks 1994	This was a within-group study of 19 midwives in the UK following a 1-day workshop on critical appraisal teaching. The outcomes measured were skills and self assessed behaviour changes. This study was excluded because it did not assess outcomes specified in the inclusion criteria and did not use objective measures of those outcomes that fulfilled our relevance criteria.
Hillson 1993	This was a within-group comparison of 29 residents who had 7 hours of lectures and journal clubs. Critical appraisal skills were assessed. This study was excluded as it did not assess outcomes specified in the inclusion criteria.
Ibbotson 1998	This was a multidisciplinary 1-day workshop teaching critical appraisal to 115 healthcare workers, and evaluating self assessed knowledge and attitudes. This study was excluded as it did not assess outcomes using objective measures.
Inui 1981	This described a seminar series for teaching critical appraisal skills to second year residents. It was excluded because no comparative data were presented.
Johnson 1995	See Reineck 1995
Kerrison 1995	This study assessed the impact of a series of 23 workshops involving an estimated 130 participants, including managers, clinicians and researchers. The main focus of the evaluation was qualitative; some quantitative data were collected. It was excluded because no comparative data were presented. Confirmation of this from the lead author was sought but not obtained.
Kitchens 1989	This was a within-group comparison and between-group cross-over trial of 83 internal medicine interns involving critical appraisal seminars and assessing knowledge. This was excluded because it had not stated the validity of instruments used.

Study	Reason for exclusion
Konen 1990	This described 5 years of experience with a curriculum teaching critical appraisal skills to family practice residents. It was excluded because no comparison group data were available, a fact confirmed by the author.
Krueger 2006	This was excluded because the intervention occurred with medical students and not qualified health workers
Landry 1994	This was a controlled trial of 146 medical students with 3 hours of large-group seminars teaching critical appraisal. Outcomes assessed were knowledge and attitudes. This was excluded as it did not teach qualified healthcare workers.
Langkamp 1992	This controlled before and after study assessed the impact of 2 didactic sessions on research design, clinical epidemiology and biostatistics followed by 8 monthly journal club sessions to 27 paediatric residents at 2 institutions. It was provisionally included pending confirmation that the intervention was critical appraisal teaching, and that it was being compared to no teaching. Unfortunately, no clarification was received and the study was effectively excluded, despite clearly being of potential relevance.
Linzer 1987a	This quasi-randomised controlled trial compared a journal club for residents teaching clinical epidemiology and biostatistical skills co-ordinated by a general medicine faculty member with a special interest and training in clinical epidemiology, biostatistics and critical appraisal, versus one co-ordinated by a chief resident. It was excluded because it did not assess the impact of teaching critical appraisal, but rather 2 alternative approaches to teaching of critical appraisal. It was identified as a potentially useful study outside the immediate context of this review.
MacAuley 1996	This study was excluded because it evaluated an instrument to aid critical appraisal and not the impact of teaching it
MacAuley 1997	This study was excluded because it evaluated an instrument to aid critical appraisal and not the impact of teaching it
MacAuley 1998	This study was excluded because it evaluated an instrument to aid critical appraisal and not the impact of teaching it
MacAuley 1999	This study was excluded because it evaluated an instrument to aid critical appraisal and not the impact of teaching it
Markert 1989	This described the development of a journal club teaching medical residents how to read and critically appraise the medical literature. It was excluded because there was no comparison group and no pre-course assessment.
Milne 1996	This study evaluated critical appraisal workshops designed to develop skills needed to make sense of evidence about effectiveness for people who give health information to the public, especially staff in consumer health information services and members of maternity self help groups. It was excluded because there was no comparison group and no pre-course assessment.
Mulvihill 1973	This study described a course in epidemiology and biostatistics to medical students. It was excluded because teaching critical appraisal did not appear to be a major component of the course. Further, there were only minimal data on evaluation.
Mulvihill 1980	See above reference by same author
Novick 1985	The second part of the epidemiology course described was concerned with the teaching of critical appraisal skills. It was excluded because there was no comparison group and no pre-course assessment.

Study	Reason for exclusion
O'Sullivan 1995	This study assessed general internal medicine residents' perception of how 2 methods of teaching critical appraisal affected their reading habits, presentation skills and critical appraisal skills. It was excluded as no comparison with no critical appraisal teaching was provided. It was identified as a potentially useful study outside the immediate context of this review.
Radack 1986	This controlled trial of 34 medical students compared 5 problem-based critical appraisal sessions with a control group. Outcomes assessed were critical appraisal skills. This was excluded as it did not teach qualified healthcare workers, and did not state the validity of instruments used.
Reineck 1995	This publication described a pilot study of a critical reading of research programme over 6 sessions for nurses following on from the preliminary work by Jean Johnson. It was excluded because there was no comparison group and no pre-course assessment, although it was indicated that the latter had been collected in the associated paper by Johnson.
Riegelman 1986	This controlled trial of 296 medical students compared critical appraisal seminars and a lecture course with control group. It was excluded as it did not teach qualified healthcare workers, and had not stated the validity of instruments used.
Romm 1989	This randomised controlled trial compared teaching critical appraisal using either small group formats or lectures. It was excluded because it had no comparison group not being taught critical appraisal. It was identified as a potentially useful study outside the immediate context of this review.
Salmi 1991	This French paper was a tutorial on the critical appraisal process and was excluded because it was not an evaluation
Sandifer 1996	This study was set in a department of public health in South Glamorgan Health Authority, UK and assessed the use of a journal club as a learning environment to practise critical appraisal skills. The proxy outcome indicators used were impact on commissioning policy and publication of letters to the editor of the journal from which the appraisers' articles were selected. It was excluded because there was no comparison group and no pre-course assessment.
Seelig 1991	This was a within-group comparison of 18 internal medicine internists following a 1-hour teaching session on critical appraisal and assessing knowledge, skills and attitudes. Ultimately this publication was excluded because it did not use a validated instrument.
Seelig 1993	This publication examined a within-group comparison of 14 internal medicine residents who had one lecture and 8 journal club sessions on critical appraisal skills. Skills, knowledge and attitudes were assessed. This was excluded as it did not state the validity of instruments used.
Stern 1995	This study was excluded as it was developing and validating an instrument to evaluate the abilities of residents to critically appraise a journal article and not the teaching of critical appraisal skills.
Viniestra 1994	By the same authors as the excluded study Alarcon 1994, this Mexican study involved the evaluation of critical appraisal skills in medical students and an attempt to relate differences observed to other characteristics, such as years of medical training. It was excluded as it did not involve teaching critical appraisal.

RCT: randomised controlled trial

Characteristics of studies awaiting assessment *[ordered by study ID]*

[McLeod 2010](#)

Methods	12 general surgery programmes were cluster-randomised to an Internet group or a moderated journal club
---------	--

McLeod 2010 (Continued)

Participants	Residents
Interventions	
Outcomes	Critical appraisal testing
Notes	

ADDITIONAL TABLES
Table 1. Summary of characteristics of excluded studies

Study author and year	Study design	Outcome measure
Alarcón 1994	Controlled before-after study	Critical appraisal skill
Bennett 1987	Controlled trial	Objectivity skills
Burls 1997	Before-after study	Self assessed knowledge and attitudes
Burstein 1996	Before-after study	Impact of structured tool in journal club
Caudill 1993	Before-after study	Knowledge and skills
Cuddy 1984	Controlled trial	Knowledge
Domholdt 1994	Comparative study	Comparison of skills of 2 groups of physical therapists
Dorsch 1990	Descriptive study	Described development of critical appraisal intervention
Fowkes 1984	Before-after study	Epidemiology knowledge after an epidemiology course with no separate data for critical appraisal
Frasca 1992	Controlled trial	Skills
Gehlbach 1980	Controlled trial	Skills
Gehlbach 1985	Controlled trial	Skills
Globerman 1993	Descriptive	Described development of course
Griffith 1988	Descriptive	Evaluation of critical appraisal course assessing satisfaction, behaviour change, attitudes, no comparison nor pre intervention assessment
Heller 1984	Controlled trial	Evaluation of general epidemiology and public health course, no data separable for critical appraisal
Heiligman 1991	Descriptive	Assessed attitudes to a journal club
Herbert 1990	Descriptive	Described critical appraisal teaching course with course evaluation
Hicks 1994	Within-group before-after study	Skills and behaviour change

Table 1. Summary of characteristics of excluded studies (Continued)

Hillson 1993	Within-group before-after study	Skills
Ibbotson 1998	Within-group before-after study	Self assessed knowledge and attitudes
Inui 1981	Descriptive	Critical appraisal teaching course description
Kerrison 1995	Within-group before-after study	Qualitative impact of workshop
Kitchens 1989	Within-group before-after comparison and between group crossover trial	Knowledge
Konen 1990	Descriptive	Described 5 years experience of teaching critical appraisal
Kreuger 2006	Randomised controlled trial	Skills appraised in medical students
Landry 1994	Controlled trial	Knowledge and attitudes
Langkamp 1992	Controlled before-after study	Assessed impact of intervention which was not clearly critical appraisal only
Linzer 1987a	Quasi-randomised controlled trial	Skills knowledge compared 2 methods of teaching critical appraisal
MacAuley 1996	Before-after study	Evaluation of tool for critical appraisal
MacAuley 1997	Before-after study	Evaluation of tool for critical appraisal
MacAuley 1998	Randomised controlled trial	Evaluation of tool for critical appraisal
MacAuley 1999	Randomised controlled trial	Evaluation of tool for critical appraisal
Markert 1989	Descriptive	Described the development of a journal club teaching critical appraisal and evaluation of this club
Milne 1996	Educational approach	Evaluated workshops teaching critical appraisal skills satisfaction
Mulvihill 1980	Descriptive	Described epidemiology course
Mulvihill 1973	Descriptive	Descriptive
Novick 1985	Evaluation of course	Evaluated critical appraisal assessed as satisfaction with course
O'Sullivan 1995	Controlled trial	Skills and behaviour assessed after 2 methods of teaching critical appraisal
Radack 1986	Controlled trial	Skills
Riegelman 1986	Controlled trial	Skills behaviour knowledge
Reineck 1995	Descriptive	Described course
Romm 1989	Randomised controlled trial	Compared teaching critical appraisal using either small group formats or lectures
Salmi 1991	Descriptive	Described the critical appraisal process

Table 1. Summary of characteristics of excluded studies (Continued)

Sandifer 1996	Evaluation of journal club	Assessed the use of a journal club as a learning environment to practise critical appraisal skills; attendance papers reviewed impact on commissioning policy letters to journals
Seelig 1991	Within-group before-after study	Skills knowledge attitudes
Seelig 1993	Within-group before-after study	Skills knowledge attitudes
Stern 1995	Validation of instrument	Developing and validating an instrument
Viniegra 1994	Controlled before-after study	Critical reading intervention and skills assessed

Table 2. Results of included trials

Study	Intervention	Knowledge outcomes	Critical appraisal	Notes
Linzer 1988	Medical journal clubs with emphasis on critical appraisal (3 to 6 hours actually attended over average 9.5 months). Between-group change assessed	Mean (SD) (change in knowledge scores post intervention) IG 1.5 (2.0), CG 0.3 (1.6); % improvement IG 25 (35), CG 0 (33) Note: dose response observed - more sessions attended = more knowledge gained	Mean (SD) scores IG 0.5 (2.2), CG 1.7 (2.2); % improvement IG 3 (20), CG 11 (16)	Benefits of randomisation may not be realised due to small numbers leading to lack of comparability of groups and openness to confounding and bias
MacRae 2004	Participants in the intervention group received 8 packages (of 2 articles) and questions designed to guide critical appraisal. Listserve discussions, a comprehensive methodologic review of the article prepared by a surgeon with training in clinical epidemiology and a clinical review were also provided.	Not reported	Overall mean score on the exam was 53.9% (+/- 9); IG 58.8% (8) versus CG 50% (8) (P < 0.001)	Not applicable
Taylor 2004	1/2 day workshop based on Critical Appraisal Skills Programme (CASP)	Total knowledge score significantly higher for the IG versus CG (ITT mean difference 2.6 (95% CI 0.6 to 4.6); explanatory analysis mean difference 3.1 (95% CI 1.1 to 5.2); IG mean 9.7 (5.3) CG 8.0 (5.1))	Ability to assess a SR: ITT mean difference 1.2 (95% CI 0.01 to 2.4) not significant. No difference was observed in the ability to appraise 'methodology' or 'relevance/generalisability' of evidence.	Trial failed to recruit the target number of individuals, thus study is underpowered. Cost data calculated at an average cost of GBP 250 per person

CG: control group
 IG: intervention group
 ITT: intention-to-treat
 SD: standard deviation
 SR: systematic review

APPENDICES

Appendix 1. Update search strategies

Ovid EMBASE, MEDLINE, PsycINFO, CINAHL (also Ebscohost)

1 education.fs.
2 ed.fs.
3 exp Education/
4 exp teaching/
5 exp Evidence-Based Medicine/
6 exp evidence based practice/
7 exp professional practice, evidence-based/
8 or/1-7
9 (critical\$ and apprais\$).ab,ti.
10 (journal\$ and club\$).ab,ti.
11 (critical\$ and read\$).ab,ti.
12 or/9-11
13 8 and 12

Ovid CDSR, CCTR, DARE, MEDLINE in process

1 (critical\$ and apprais\$).ab,ti.
2 (journal\$ and club\$).ab,ti.
3 (critical\$ and read\$).ab,ti.
4 or/1-3

ISI Web of Knowledge

Science Citation Index Expanded, Social Sciences Citation Index

#1 TS=("critical* apprais*") Timespan=2007-2008
#2 TS=("journal* club*") Timespan=2007-2008
#3 TS=("critical* read*") Timespan=2007-2008
#4 #3 OR #2 OR #1

Proquest Database: LISA

KW=(teach* AND critical* AND apprais*)
or
DE=(evidence-based medicine) and KW=(apprais* or teach* or train*)

Proquest Database: ERIC

KW=(teach* OR educat* or learn*) AND critical* AND apprais*

EPOC Specialised Register

"critical* apprais*" or "journal club" or "book club" or "apprais*

WHAT'S NEW

Date	Event	Description
29 September 2011	New citation required but conclusions have not changed	Three new authors; two new studies added.
29 September 2010	New search has been performed	Review updated

HISTORY

Protocol first published: Issue 4, 1998

Review first published: Issue 3, 2001

Date	Event	Description
26 August 2008	Amended	Converted to new review format.
14 May 2001	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

First version of the review: Julie Parkes wrote the protocol, performed the searching, read the abstracts and retrieved and appraised relevant evidence. Jon Deeks acted as senior supervisor to these processes. Ruairidh Milne, Jon Deeks and Chris Hyde selected trials for inclusion in the review. Two researchers (JP, CH) extracted the necessary data using a checklist developed by EPOC, modified and amended for the purposes of this review. Quantitative data from each study were reported in natural units in the results table. Chris Hyde and Julie Parkes wrote the first draft of the review. Chris Hyde, Jon Deeks and Ruairidh Milne were involved in the discussion.

Updated version of the review: **Tanya Horsley:** review supervision, title and abstract and full-text relevance assessment, data extraction, risk of bias assessments, interpretation and summary of data (text and tables), initial draft and critical revision of manuscript. **Chris Hyde:** title and abstract screening, full-text retrieval, critical revision of the manuscript. **Nancy Santesso:** review and revision (where applicable) of original searches, initial review supervision, search execution, initial relevance assessment and de-duplication, title and abstract and full-text relevance assessment, critical revision of the final manuscript. **Ruth Stewart:** title and abstract and full-text relevance assessment, critical revision of the final manuscript. **Julie Parkes and Ruairidh Milne** were original review authors and provided comment on the final draft.

DECLARATIONS OF INTEREST

First version of the review: from 1993 to 1996 Dr Ruairidh Milne worked with the Oxford-based Critical Appraisal Skills Programme. He has continued to teach critical appraisal and is occasionally paid for doing so.

Updated version of the review: T. Horsley - none to declare, C. Hyde - none to declare, N. Santesso - none to declare, R. Stewart - Dr. Stewart's work at the EPPI-Centre involves occasionally teaching critical appraisal skills and compensation is provided. Authors with continuing authorship: J. Parkes - none to declare, R. Milne - none to declare.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- First review: National R&D Programme Evaluating Methods to Promote the Implementation of Research, UK.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

See search string modification ([Search methods for identification of studies](#)).

INDEX TERMS

Medical Subject Headings (MeSH)

*Awareness; *Comprehension; *Information Literacy; Delivery of Health Care; Health Personnel [*education]; Translational Medical Research

MeSH check words

Humans