Langenbeck's Arch Surg (1999) 384:409–416

Gerd Antes
Daniel Galandi
Bertil Bouillon

What is evidence-based medicine?

Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research (DL Sackett).

Received: 9 June 1999
Accepted: 27 August 1999

G. Antes (✉) · D. Galandi
German Cochrane Centre,
Department of Medical Informatics,
University Hospital of Freiburg,
Stefan-Meier-Str. 26, D-79104 Freiburg,
Germany
e-mail: mail@cochrane.de
Tel.: +49-761-2037615
Fax: +49-761-2036712

B. Bouillon
Surgical Clinic Merheim,
2nd Department of Surgery,
University of Cologne, Cologne, Germany

Abstract *Introduction:* Three challenges that physicians and decision makers in the health care systems have to meet are a remarkable proportion of medical decisions without a sufficient base of scientific evidence, a slow and opaque process of integrating scientific knowledge into medical practice and a steadily decreasing half-life period of the medical knowledge. *Discussion:* During the last two decades, a number of projects have faced these problems and forced the development of evidence-based medicine (EBM). This concept claims the explicit conscientious use of current evidence from clinical research combined with the personal expertise in the process of medical decision making. The following article explains the main steps of practising and teaching EBM illustrated by a clinical example.

Key words Evidence-based medicine · Clinical study · Clinical scenario · Teaching EBM

Introduction

These sentences, phrased in a frequently cited editorial [1], characterise the principles of evidence-based medicine (EBM) as it is understood and propagated by many physicians and researchers as well as institutions and organisations of the health systems. The discussion regarding the need of a better integration of the results of medical research is mainly provoked by three actualities: (1) the percentage of medical decisions based on scientific evidence is estimated as 15–40%, sometimes even lower [2]; (2) there is an enormous time lag between assured scientific knowledge and the introduction into medical routine, empirically shown to be 8–10 years [3]; and (3) the medical knowledge is, at rapidly increasing speed, becoming out of date, with a steadily decreasing half-life period of about 5–45 years today, depending on speciality [4, 5].

The implications of these statements have been debated since 1970 [6], leading to a series of projects and developments in North America and the UK over the last 20 years. Stimulated by the crisis of their health care systems and other factors, many central European countries began to take part in the discussion a few years ago. Because of its broad dissemination, there is a need for all medical specialities to get involved in this development and to consider the impact EBM already has and will have for them in future.

The concept of EBM

The global target of EBM is to improve health care by bridging the gap between research and clinical practice, thus allowing medical decisions to be based on the best available scientific evidence [7, 8]. To achieve this target, the available evidence or knowledge has to be accessed in a formalised and operationalised way for an optimal output and benefit. The methodological framework for an EBM is organised in the following four steps:

1. Translating a clinical problem into a question that can be answered

2. Searching for the best evidence
3. Critically appraising the evidence
4. Applying the identified and appraised information in clinical practice

At the end of the process, a fifth step can be added, which is the evaluation of the whole process and of the clinical performance.

Translating a clinical problem into a question that can be answered

Use of everyday language is not sufficient for asking questions that will be answered by searching the medical literature via bibliographic databases. Questions suitable for a systematic search have to follow a certain logic that breaks down the overall question, transferring it into an operationalised form. Well-constructed clinical questions should contain four elements [7]:

1. The patient and/or problem addressed.
2. The intervention we are interested in (e.g. a diagnostic test, a treatment or a prognostic factor).
3. A comparison intervention, if relevant.
4. The clinical outcome or outcomes we and the patient are interested in.

For diagnostic decisions, a well-formulated question should enable the identification of an appropriate diagnostic test with its characteristics (sensitivity and specificity or likelihood ratios). These test properties are essential for evaluating a diagnostic test and its benefit for the clinical decision-making process. With the knowledge of the patients' probability of being ill (prevalence or pre-test probability), the characteristics of a diagnostic test and the result of this test in a patient, we can calculate the post-test probability and assess the diagnostic result.

Searching for the best evidence

To identify the available evidence, modern tools such as electronic databases have to be used [9]. Today, this usually means using Medline, mainly because the access is free. However, it should be noted that even a qualified Medline-search may not produce the best existing evidence [10]. First, the reason may lie in the difficult index structure, which can be a serious obstacle for a successful search, thus inhibiting finding studies contained in Medline. Second, aside from the complex index structure, a remarkable number of articles are indexed incorrectly or incompletely [11]. Third, Medline does not comprise all journals, leaving out some journals with a large number of studies, or even worse, not covering a number of special areas. Especially in Europe, the medical database Embase (Elsevier) may be a better choice, particularly for pharmaceutical studies. New articles are included faster into Embase than into Medline and it contains journals that are not covered by Medline. Regardless of the selected databases, the search for evidence must be highly specific. In most cases, the first appropriate article has to be used because of time constraints and the conditions of everyday clinical practice, which rarely allow use of information from several articles.

To overcome this limitation and provide the busy clinicians and decision makers in the health care systems with reliable information, synthesising the existing research results is an essential step beyond relying on single studies only [12, 13]. The Cochrane Collaboration (CC), an international network of physicians, scientists, health care institutions and patients [14] started to prepare, maintain and disseminate systematic reviews of all types of health care interventions. Systematic reviews are produced under conditions designed to overcome the intrinsic errors of classical narrative reviews [12]. During recent years, they have been acknowledged as the most reliable source for appraising the effectiveness of an intervention. Therefore, in the classification of reliability of evidence, systematic reviews that are based on high-quality randomised controlled trials (RCT) receive evidence level 1 (Table 1) [15]. Logically, a search should start with systematic reviews, going down the evidence hierarchy of the table if there is no review. An easy-to-use database is the Cochrane Library [16], which contains 628 systematic reviews and 593 protocols of reviews (Issue III/1999). The Cochrane Library also contains a literature database with all clinical study citations from Medline and an increasing number from Embase. These citations are complemented by a large number of citations which were identified by systematically hand searching journals not covered by electronic databases. Providing these data on clinical studies makes the Cochrane Library one of the most

Table 1 Levels of evidence [15]

Level	Type of investigation
Ia	Evidence obtained from meta-analysis of randomised controlled trials
Ib	Evidence obtained from at least one randomised controlled trial
IIa	Evidence obtained from at least one well-designed controlled study without randomisation
IIb	Evidence obtained from at least one other type of well-designed quasi-experimental study
III	Evidence obtained from well-designed non-experimental studies, such as comparative studies, correlational studies, and case studies
IV	Evidence obtained from expert committee reports or opinions and/or clinical experiences of respected authorities

comprehensive if not the best single source of evidence [17]. A list of Cochrane reviews dealing with surgical topics is maintained by the Royal College of Surgeons on the internet under http://www.rcseng.ac.uk/public/infores/reso_ir.htm.

An alternative approach to reduce the workload is the use of condensed versions of original articles provided by the secondary literature. Well-known examples are the journal *Evidence-Based Medicine* [18] or its electronic version *Best Evidence* which comprises EBM and the ACP journal club and appears on CD-ROM. Original articles are identified by systematically screening a large number of high-quality journals. The articles, selected according to methodological quality and clinical relevance, are reduced to an abstract under a fixed format which is published together with an independent comment.

Critically appraising the evidence

The identified scientific evidence has to be critically appraised. This step is absolutely essential, because a large proportion of published medical research lacks methodological rigour to be reliable. Therefore, a number of structured and fairly simple guides have been developed to enable people without research expertise to evaluate clinical articles and use them as base for their decision making. A widely used series of articles are the *Users' Guides to the Medical Literature* published between 1993 and 1998 in the *Journal of the American Medical Association* (a detailed reference list can be found under http://www.cochrane.de/deutsch). A checklist may be used to assess articles with respect to several topics, e.g. randomisation, intention-to-treat analysis, discussion of benefit versus harm and costs etc., in an article about treatment.

Another approach to judge reliability of information is using levels of evidence (Table 1) that express the validity of the evidence according to the underlying study design [15]. There has been a long, still ongoing debate on the quality assessment of clinical trials. Various scales of different length and complexity exist [19] but, so far, no generally agreed method has been derived. While many of the scales are very complex or even too complex for extensive practical use, recently, a few components have been identified as being particularly important. The dominant factors are randomisation, blinding and dealing with drop-outs [20]. If studies are open because blinding is not possible or is not done, concealment of treatment allocation has been shown to play a crucial role. Lack of concealment may lead to serious bias, normally an overestimation of the treatment effect [21].

An unsolvable problem is the serious confounding of conducting and reporting trials [22]. From the articles it often cannot be concluded whether crucial quality requirements in a study were neglected or only left out in the report. The CONSORT-Statement, a guideline for reporting clinical studies, was developed to lead to more standardisation in study reports and is now adhered by more than 70 journals [23]. Work on a similar type of guideline for meta-analyses is still in progress. Both statements together should gradually lead to an improvement if they become part of the instructions for authors by a majority of the biomedical journals.

Applying the identified and appraised information to clinical practice

Successfully working through the previous steps produces reliable information which has to be applied to the specific, individual patient. As mentioned above, applying a diagnostic test with known parameters allows calculation of the post-test probability of being ill. The implications of a potential false positive or false negative classification have to be assessed with respect to the particular disease and the conditions of the patient [7, 24, 25].

Information about therapeutic decisions can be presented in different ways. Widely used, especially in information from the pharmaceutical industry, is the relative risk reduction (RRR), which is the proportional reduction of the risk to suffer a negative event between the intervention and the control group. Consider a patient who has a risk of suffering a certain negative event (e.g. stroke during the next year because of carotid artery stenosis) of about 30% without any treatment. Treating the patient (e.g. carotid endarterectomy) would reduce this risk to 15%. We have an RRR of (30%–15%)/15%=50%. This clearly seems to be worth the effort. Suppose the risk of the patient to suffer stroke in the next year would only be 0.3% and treating the patient would reduce this to 0.15%. In this example, the RRR is still (0.3%–0.15%)/0.15%=50%, but the therapy lacks clinical relevance because the baseline risk is very low and not worth taking the risk of the therapeutic intervention. Thus, the RRR fails to visualise this important difference.

A better approach to express the effects of therapeutic interventions is the absolute risk reduction (ARR), which is the *difference* between the risk to suffer a certain event in the control group and this risk in the intervention group. In the first example, we have an ARR of 30%–15%=15%, which again recommends treatment of the patient. In the second example, the ARR is only 0.3%–0.15%=0.15%, which is not worth thinking about.

To put the information of the ARR in a more illustrative figure, we can use the inverse of the ARR, which is the number needed to treat (NNT). The NNT describes

the number of necessary therapeutic interventions to allow one additional success by using the considered therapy. In the first example, the NNT is 1/0.15=7. Seven patients have to undergo carotid endarterectomy to prevent one stroke during the next year. In the second example, the NNT would be 1/0.0015=667 and the irrelevance of the therapeutic intervention in this example becomes obvious. The NNT is an intuitively applicable tool for making decisions about therapeutic interventions and integrating the patients' preferences in the decision-making process. Of course the physician is forced to define a kind of threshold-NNT, depending on the individual clinical situation and the relationship between harm and benefit caused by a therapeutic procedure under the particular circumstances.

Regardless of how the information is presented in an article, some basic figures, such as the number of patients in the treatment groups and the number of adverse outcomes in each group, is enough to calculate ARR and NNT. Beneath these technical aspects, the individual situation of the patient and his expectations have to be taken into account in this step [7, 26, 27].

Practicing and teaching EBM: a clinical scenario

Consider the following situation. During ward rounds in the outpatient department, students saw a 25-year-old patient who twisted his left knee during a soccer game the day before. He complained about pain and reduced extension and flexion of his left knee. The students examined the patient's knee and recognised swelling, pain and a reduced function. The medial and lateral ligaments were stable. They were not quite sure about the diagnosis. They suspected a distorsion of the knee, but one student suggested that this patient could also have a rupture of the anterior cruciate ligament (ACL). Following the four steps of EBM, the students translated the clinical problem into an answerable question, which can serve as a start for a literature search.

Is there a sensitive, easy-to-apply clinical examination method to prove ACL rupture in a 25 year-old patient with pain and reduced function of the knee after a sports injury?

After a brief instruction on how to do a short literature review in the KnowledgeFinder [28], using "diagnosis of ACL rupture", they found two comparative studies that could answer their question (complete reference see reference list [29, 30]):

Liu S.H. et al. The diagnosis of acute complete tears of the anterior cruciate ligament. Comparison of MRI, arthrometry and clinical examination

Rose N.E. et al. A comparison of accuracy between clinical examination and magnetic resonance imaging in the diagnosis of meniscal and anterior cruciate ligament tears

Both articles make contributions to the interesting clinical problem, but the trial of Liu et al. has only 38 participants compared with 154 in the trial published by Rose et al. Because of the greater number of patients included in the study, the students decided to get a copy of the latter investigation to make a critical appraisal of the identified evidence. According to the instructions given in the *Users' Guides to the Medical Literature: 'How to use an article about a diagnostic test?'*, they answered the questions detailed in Appendix I [24, 25]:

1. Are the results of the study valid?
The comparison with the reference standard was not blind and independent because it was applied by the same person, but the reference standard (arthroscopy) was applied independently from the result of the test under evaluation (clinical examination). The patient sample seems to cover the actual clinical situation and the process was described in sufficient detail. The students felt a little uneasy about the lack of an independent comparison with the reference standard but after a short discussion they decided that this piece of evidence is valid.

2. What are the results?
The authors do not present likelihood ratios, but the given information with specificity of 99% (95% confidence interval 93.8–99.9) and a sensitivity of 100% (75.2–100) of the Lachmann test to detect an ACL rupture is sufficient. In this situation, a detailed estimation of the post-test probability by combining prevalence and likelihood ratios is not essential [24, 25, 31].

3. Will the results help me in caring for my patient?
The students agreed that the information provided by this article was applicable to their patient and was very helpful to decide about the right diagnosis.

Based on these findings, there seems to be enough evidence that the Lachman test as a clinical test is sufficiently sensitive to detect ACL ruptures. After demonstration of the Lachman test by the surgeon and examination of the patient, it became clear that the Lachman test was positive. The students concluded that the patient had an ACL rupture.

Faced with this diagnosis, the patient asked which consequences this injury had for him, especially for his sport activities, and wanted to get more information concerning the best therapy in his situation – in particular the advantages and disadvantages of surgical treatment.

Remembering the concepts of EBM, the students formulated the following question.

In a 25-year-old patient with ACL rupture, does surgical treatment lead to better results than conservative treatment with regard to sport activities?

They conducted a quick search of the literature beginning with the Cochrane Library and found only a protocol for a Cochrane review [32] but no completed systematic review on this topic. Not being successful on the highest level of evidence, the students searched the KnowledgeFinder Medline-System with the string "treatment of ACL rupture" and the search engine came up with the following citations (for complete reference, see reference list [33, 34, 35, 36]):

Rauch G. et al. Is conservative treatment of partial or complete anterior cruciate ligament rupture still justified?

Anderson C. et al. Knee function after surgical or non-surgical treatment of acute rupture of the anterior cruciate ligament: a randomized study with long-term follow-up period

Daniel D.M. et al. Fate of the ACL injured patient. A prospective outcome study

Fink C. et al. Long-term outcome of conservative or surgical therapy of anterior cruciate ligament rupture

In summary, the abstracts of these articles showed that both surgical and conservative therapy are still an alternative worth considering.

For a more detailed appraisal, the students took a paper copy of the complete article from Anderson et al., following the guides in Appendix II coming from the *Users' Guides to the Medical Literature 'How to use an article about therapy'* [26, 27].

1. Are the results of the study valid?

Concerning the validity of the trial, they concluded that this article reports a RCT, all patients who entered the trial were included in the final evaluation except for the properly explained withdrawals and all patients were analysed in the groups to which they had been randomised. The treatment groups were similar at the beginning of the trial, the groups were treated equally aside from the experimental treatment and the follow-up period was long enough. Blinding was not possible because of the comparison of surgical and conservative treatment.

After a short discussion, the students agreed that the trial was valid and the methodological aspects gave no obvious reason to suspect substantial bias within the trial.

2. What are the results?

With respect to sports activities, the authors explain that 69% of the conservatively treated patients did not return to competitive sports at the follow-up examination compared with 50% in the ACL repair group (data are pooled from different techniques of ACL repair). The distribution of patients who took part in competitive sports before the injury did not differ among the groups. This result indicates an ARR of 50%–31%=19% between conservative treatment and surgical treatment. To put this in a more illustrative way, the surgeon demonstrated how to calculate the NNT, which is 1/ARR or 1/0.19=5. In this case, if only five patients undergo surgical repair of the ACL, one more patient will be able to return to competitive sports compared with conservative treatment. The article does not give any information about confidence intervals, but a P value of lower than 0.05 is mentioned so the students believe this result to be significant.

3. Will the results help me in caring for my patients?

The students concluded that the results can be applied to their patient, that the trial gives sufficient information about the main clinical outcomes and that the expected treatment benefit is worth the potential harms and costs, especially in a young patient who wants to continue his sporting activities. In a different clinical situation with an older patient who was working in an office without any sportive ambitions and without subjective knee instability, the conservative treatment would probably be the better choice.

The patient confirmed that soccer was one of his favourite sports that he wanted to continue and agreed to a surgical reconstruction of his ACL after the surgeon's explanation of the treatment alternatives and the advantages of a surgical approach, especially with respect to returning to sporting activities with an NNT of about 5.

Perspectives

EBM is the process of systematically finding, appraising and using contemporaneous research findings as the basis for clinical decision making. However, searching for and using published literature is heavily influenced by many sources of bias which are inherent in the research and publication process.

Particularly serious is the publication bias, which means the larger probability of clinical trials with significant comparisons to be published [37]. For several reasons, related to the attitude and behaviour of researchers, publishers and sponsors, significant studies are published more often and faster than trials that do not find significant differences between treatment groups [38]. Consequently, averaging over several studies in a review can lead to an overestimation of the summarised treatment effect. The solution for this problem would be the preregistration of all clinical trials with compulsory publication [39], regardless of outcome and completion. Implementing such a register would be possible under the existing logistic conditions but seems nonetheless far from realisation.

Finding published studies is a serious problem, described as retrieval bias. Incomplete databases and imperfect indexing are the main reasons for this phenomenon [10]. The comprehensive and unbiased identification of trials is further complicated by the language bias. The tendency to publish significant trials from non-English language areas in English-language journals combined with a better representation of these journals in libraries and databases leads to a systematic over-representation of trials that demonstrate large (significant) treatment effects [40]. The common practice of concentrating on publications in English is not supported by the quality of the trials in general. Empirical research proved that there are no differences with respect to trial quality between English- and non-English-language journals [41]. Finally, personal career interests and the influence of impact factors additionally complicate the path from producing clinical evidence to its use in medical practice.

The concepts described above and illustrated by the clinical example follow the classical EBM as described by Sackett and others [7]. Developed under ideal situations, this approach seems to be applied in practical conditions only with considerable difficulties. The skills to integrate these steps into everyday practice, under time-constraints and lack of logistic support, are a major obstacle. Introducing these decision structures would ask, at least in the beginning, for additional resources which are hardly available these days.

The surgeon who wants to practice EBM is faced with the problem that a considerable proportion of surgical treatments cannot be evaluated in RCTs, which are recommended by the protagonists of EBM [42]. The lack of controlled trials is even more serious in the comparison of surgical and conservative treatments for several reasons, such as patients' general preferences of surgical treatment. In absence of systematic reviews or RCTs, the levels of evidence should be used as a kind of guideline to base clinical decisions on pieces of evidence of a lower level [15].

The classical EBM concentrates on the individual patient and is heavily criticised for this narrow focus [43]. The EBM approach has been expanded to include the view to populations and their optimal care, coining the expression evidence-based health care (EBHC) [44]. This broader view is necessary and becomes increasingly more important under the present discussion of rationing medical care.

Instead of relying on original literature, the use of secondary literature (e.g. journals like EBM or systematic reviews, as produced by the CC) may offer a better, more realistic choice. Especially the CC, in a tremendous worldwide effort, aims to prepare, maintain and disseminate systematic reviews on the effects of health care, which in near future should offer a knowledge base allowing a major step forward towards a science-based medical decision making.

The deficiencies of the transfer process of scientific medical knowledge into medical practice facilitate fairly homogeneous decisions by single practitioners or in single hospitals, but lead to considerable between-practitioners or between-hospital variation, resulting in different decisions under absolutely identical conditions. Basing the everyday decision making on scientific, readily available knowledge should reduce the variability and thus lead to an improved health care.

These statements have to be seen in connection with the growing information crisis, with an exponentially growing knowledge and the transition from publications on paper to electronic media, largely based on the internet [45]. Huge changes have to be expected in the future, presenting medical knowledge to the public and to patients (the e-patient, the electronically informed-patient), thus leading to absolutely new challenges for the medical profession. The concepts of EBM provide helpful tools to cope with these new challenges.

Acknowledgements We thank Stefan Sauerland, Biochemical and Experimental Division, 2nd Department of Surgery, University of Cologne, Cologne, Germany, for his help in searching and analysing the literature.

Appendix I [24, 25]

1. Are the results of the study valid?

Primary guides

Was there an independent, blind comparison with a reference standard?

Did the patient sample include an appropriate spectrum of patients to whom the diagnostic test will be applied in clinical practice?

Secondary guides

Did the results of the test being evaluated influence the decision to perform the reference standard?

Were the methods for performing the test described in sufficient detail to permit replication?

2. What are the results?

Are likelihood ratios for the test results presented or data necessary for their calculation provided?

3. Will the results help me in caring for my patients?

Will the reproducibility of the test result and its interpretation be satisfactory in my setting?

Are the results applicable to my patient?
Will the results change my management?
Will patients be better of as a result of the test?

Appendix II [26, 27]

1. Are the results of the study valid?

Primary guides

Was the assignment of patients to treatment groups randomised?

Were all patients who entered the trial properly accounted for and attributed at its conclusion?

Was follow-up complete?

Were all patients analysed in the groups to which they were randomised?

Secondary guides

Were all patients, health workers and study personnel blinded to treatment?

Were the groups similar at the start of the trial?

Aside from the experimental intervention, were the groups treated equally?

2. What are the results?

How large was the treatment effect?
How precise was the estimate of the treatment effect?

3. Will the results help me in caring for my patients?

Can the results be applied to my patient care?

Were all clinically important outcomes considered?

Are the likely treatment benefits worth the potential harms and costs?

References

1. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS (1996) Evidence based medicine: what it is and what it isn't. BMJ 312:71–72
2. Ellis J, Mulligan I, Rowe J, Sackett DL (1995) Inpatient general medicine is evidence based. Lancet 346:407–410
3. Antmann EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC (1992) A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. JAMA 268:240–248
4. Ramsey PG, Carline JD, Inui TS, Larson EB, LoGerfo JP, Norcini JJ, Wenrich MD (1991) Changes over time in the knowledge base of practicing internists. JAMA 266:1103–1107
5. Hall JC, Platell C (1997) Half-life truth in surgical literature. Lancet 350:1752
6. Cochrane AL (1972) Effictiveness and efficiency. Random reflections on health services. Nuffield Provincial Hospitals Trust, London
7. Sackett DL, Richardson WS, Rosenberg WMC, Haynes RB (1997) Evidence-based medicine. How to practice and teach EBM. Churchill Livingstone, Edinburgh
8. Antes G (1998) Evidence-based medicine. Internist 39:899–908
9. Antes G (1997) EBM praktizieren – Wie finde ich die Fragen auf meine Fragen? Münchener Medizinische Wochenschrift 139:685–688
10. Dickersin K, Scherer R, Lefebvre C (1994) Identifying relevant studies for systematic reviews. BMJ 309:1286–1291
11. Perleth M, Antes G (1998) Evidenzbasierte Medizin. MMV Verlag, München
12. Mulrow CD (1987) The medical review article: state of the science. Ann Intern Med 106:485–488
13. Antes G, Bassler D, Galandi D (1999) Systematische Übersichtsarbeiten. Ihre Rolle in einer Evidenz-basierten Gesundheitsversorgung. Dtsch Ärztebl 96:A616–A622
14. Mulrow C, Oxman AD (1997) The Cochrane handbook. In: The Cochrane Library, Update Software, Oxford
15. AHCPR (1992) Acute pain management: operative or medical procedures and trauma. Clinical practice guideline No. 1. AHCPR Publication No. 92–0032
16. The Cochrane Library (quaterly updated). Update Software, Oxford
17. Egger M, Smith DS (1998) Bias in location and selection of studies. BMJ 316:61–66
18. Evidence-based medicine. BMJ Publishing Group, Journals Marketing Dept. PO Box 299 London WC1H 9TD
19. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S (1995) Assessing the quality of randomized controlled clinical trials: an annotated bibliography of scales and checklists. Control Clin Trials 16:62–73
20. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, NcQuay HJ (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 17:1–12
21. Schulz KF, Chalmers I, Hayes RJ, Altman DG (1995) Empirical evidence of bias. JAMA 273:408–412
22. Liberati A, Himel HN, Chalmers TC (1986) A quality assessment of randomized control trials of primary treatment of breast cancer. J Clin Oncol 4:942–951
23. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Schulz KF, Simel D, Stroup DF (1996) Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 276:637–649
24. Jaeschke R, Guyatt G, Sackett DL (1994) Users' guides to the medical literature. III. How to use an article about a diagnostic test A. Are the results of the study valid? JAMA 271:389–391
25. Jaeschke R, Guyatt G, Sackett DL (1994) Users' guides to the medical literature. III. How to use an article about a diagnostic test B. What are the results and will they help me in caring for my patients? JAMA 271:703–707
26. Guyatt G, Sackett DL, Cook DJ (1993) Users' guides to the medical literature. II. How to use an article about therapy or prevention A. Are the results of the study valid? JAMA 270:2598–2601
27. Guyatt G, Sackett DL, Cook DJ (1994) Users' guides to the medical literature. II. How to use an article about therapy or prevention B. What are the results and will they help me in caring for my patients? JAMA 271:59–63
28. KnowledgeFinder, Aries Systems Corporation, 200 Sutton Street, North Andover, MA 01845 USA

29. Liu SH, Osti L, Henry M, Bocchi L (1995) The diagnosis of acute complete tears of the anterior cruciate ligament. Comparison of MRI, arthrometry and clinical examination. J Bone Joint Surg Br 77:586–588
30. Rose NE, Gold SM (1996) A comparison of accuracy between clinical examination and magnetic resonance imaging in the diagnosis of meniscal and anterior cruciate ligament tears. Arthroscopy 12:398–405
31. Sackett DL, Haynes RB, Tugwell P (1991) Clinical epidemiology. A basic science for clinical medicine. Littleton, Brown Boston Mass
32. Linko E, Harilainen A, Malmivaara A, Seitsalo S (1999) Operative treatment for anterior cruciate ligament ruptures in adults (protocol for a Cochrane review). In: The Cochrane Library, issue 1. Update Software, Oxford
33. Rauch G, Wirth T, Dorner P, Griss P (1991) Is conservative treatment of partial or complete anterior cruciate ligament rupture still justified? Z Orthop Ihre Grenzgeb 129:438–446
34. Anderson C, Odensten M, Gillquist J (1991) Knee function after surgical or nonsurgical treatment of acute rupture of the anterior cruciate ligament: a randomized study with long term follow-up period. Clin Orthop 264:255–263
35. Daniel DM, Stone ML, Dobson BE, Tithian DC, Rossman DJ, Kaufmann KR (1994) Fate of the ACL injured patient. A prospective outcome study. Am J Sports Med 22:632–644
36. Fink C, Hoser C, Benedetto KP, Hackl W, Gabi M (1996) Long term outcome of conservative or surgical therapy of anterior crusciate ligament rupture. Unfallchirurg 99:964–969
37. Dickersin K (1990) The existence of publication bias and risk factors for its occurrence. JAMA 263:1385–1389
38. Simes RJ (1987) Confronting publication bias: a cohort design for meta-analysis. Stat Med 6:11–29
39. Simes RJ (1986) Publication bias: the case for an international registry of clinical trials. J Clin Oncol 4:1529–1541
40. Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G (1997) Language bias in randomised controlled trials published in English and German. Lancet 350:326–329
41. Moher D, Fortin P, Jadad AR, Jüni P, Klassen T, LeLorrier J, Liberati A, Linde K, Penna A (1996) Completeness of reporting of trials published in languages other than English: implications for the conduct of systematic reviews. Lancet 347:363–366
42. Lefering R, Neugebauer E (1998) Problems of randomized controlled trials (RCT) in surgery. In: Abel U, Koch A (eds) Nonrandomized comparative clinical studies. Symposium Publishing, Düsseldorf, pp 67–75
43. Maynard A (1997) Evidence-based medicine: an incomplete method for informing treatment choices. Lancet 349:126–128
44. Muir Gray JA (1997) Evidence-based healthcare. Churchill Linvingstone, New York
45. Olkin I (1995) Statistical and theoretical considerations in meta-analysis. J Clin Epidemiol 48:133–146