The pros and cons of evidence-based medicine

Peter Croft [1], Antti Malmivaara [2], Maurits van Tulder [3]

[1] Professor of Primary Care Epidemiology, Arthritis Research UK National Primary Care Centre, Keele University, Keele, Staffordshire UK
p.r.croft@cphc.keele.ac.uk
[2] Senior Medical Officer, Centre for Health and Social Economics, National Institute for Health and Welfare, Helsinki, Finland
antti.malmivaara@thl.fi
[3] Professor of Health Technology Assessment, Department of Health Sciences and the EMGO Institute for Health and Care Research, VU University, Amsterdam, the Netherlands
maurits.van.tulder@falw.vu.nl

Abstract
Evidence-based medicine has evolved with at least three distinct components: the core evidence, the systematic methods of reviewing and synthesising the evidence, and the use and application of the evidence in practice. This paper considers the tensions that exist because of different perceptions about the value, quality and interpretation of clinical and research evidence. The current challenges lie in developing better means of integrating different types of evidence from a range of sources to improve decisions about individual patients, and embedding guidelines in personal and stimulating clinical practice.

Introduction
In a debate “Evidence Based Medicine: Saviour or Pariah” at the 9th International Forum on Low Back Pain Research in Primary Care, evidence based medicine (EBM) made it as saviour but it was close. Championed by Sackett in the 1990’s (1), EBM has received both support and criticism from researchers, clinicians and policy makers. In this paper, we consider some common arguments for and against EBM.

What exactly is EBM?
“Evidence based medicine” conveys the idea that up-to-date evidence can be used and applied consistently in clinical practice, in combination with the clinician’s individual expertise and the patient’s own preferences and expectations, to achieve the best possible outcomes. Evidence is provided by original research and by systematic reviews that find, select, judge, and integrate the evidence on any topic. This provides the basis for treatment guidelines, which incorporate expert interpretation of available evidence whilst taking additional arguments into account, such as treatment availability, costs and ethical aspects.
Hence, we move from evidence to practice.

Debate
Experience- or evidence-based medicine?
Clinical proponents of EBM have emphasised the range of evidence that can be used in clinical decision-making (2,3). But “evidence” can apply to any observation, so proponents could argue that there is always evidence, even if it is clinical experience without research data (4). The acronym may in reality mean "experience-based medicine". Much of the evidence in any clinical guideline is likely to come from expert opinion. Most guideline committees use expert consensus to reach decisions, drawing on a mixture of available scientific research and clinical practice and observation, with clinical judgement still playing a key role. The risk of bias may be considerable, and the term “evidence-based clinical guidelines” misleading.

Clinical practice has long been dominated by expert opinion, and there are many examples where common interventions propounded by expert opinion were ineffective or harmful. The increase in number of randomised controlled trials (RCTs) and systematic reviews, and uptake of such evidence in guidelines, has made clinical practice more transparent, reduced variation, and diminished use of harmful interventions (5,6). Initiatives like the Cochrane Collaboration and Consort Statements have encouraged trial registration and better quality trials, increasing the amount and quality of research evidence available for guidelines and decision making (7,8). Many studies now focus on developing and evaluating optimal strategies for guideline implementation in the real world of practice (9,10).

Some research evidence is better than none, and presents an advance on expert opinion. It has resulted in better, more realistic treatments. Bed rest was the mainstay of back pain treatment until RCTs showed that staying active is more beneficial (11,12). With no evidence, practice may not have changed and physicians might still be advocating bed rest. Of course, observant clinicians had long promoted exercise and critically questioned the need for rest, but it helped when trials provided independent uncluttered results.

Strong focus on RCTs

EBM has been accused of canonising the RCT, and dogmatically refusing to acknowledge other sources of valid data about the outcome of interventions (13,14). The focus on RCTs as the only design paradigm to address clinical questions is wrong. RCTs are the gold standard for questions related to the effectiveness of preventive or therapeutic interventions, but not necessarily for questions related to adverse effects, prognosis or diagnosis. Evidence on effectiveness from non-randomised observational studies may be utilized if RCT evidence is unavailable. A new method of grading evidence (GRADE (15,16)), has been recommended by the Cochrane Collaboration (17). In this, RCT evidence starts as ‘high quality evidence’, but may be lowered (even to ‘very low quality evidence’) if there are deficiencies in validity or precision, inconsistencies in outcomes, indirectness of evidence, or publication bias. Observational evidence starts as ‘low quality evidence’, but may rise to the highest level if there are factors strengthening the evidence (strong association, dose-response effect, or confounding that underestimates effectiveness), and no factors diluting or reducing it (18).

Practical or ethical reasons may prevent assessment of effectiveness by RCTs or non-randomised trials. Before-after experiments are sometimes valued in EBM if treatment has a dramatic effect and the natural course of the condition is stable.
Interventions accepted without RCT evidence include insulin for diabetes, suturing for large wounds, and defibrillation for ventricular fibrillation (19). As EBM moves to evaluate research implementation, strong observational alternatives are needed to counter the impossibility of experimentally randomising every component of interventions and translating the results into everyday practice.

**Clinical practice is more complex than trials**

If we accept that EBM is useful and that practice should be supported by adequate research evidence, its implementation in clinical practice is a major challenge. Clinical guidelines are advocated as a means to disseminate evidence and support clinical decisions, but this needs active support. Implementation of guidelines is difficult (20). General practitioners in the Netherlands have more than 90 clinical guidelines so that many neither read nor use them all. German general practitioners know about and agree with the content of back pain guidelines, but view an excessive number of information sources as a barrier to engagement with guidelines in practice (21). Using strategies that address barriers to change may be crucial to increase the usefulness of EBM.

Complexity is not limited to implementation, but lies in the very nature of clinical practice. Consider a patient who visits his doctor after months of low back pain mixed with work absence and developing depression. Cochrane collaboration reviews evaluate single interventions (22), so the doctor needs to be familiar with the results of all relevant reviews on treatments for low back pain and depression, in primary care as well as occupational settings. In practice, by contrast, patients will often be offered combined interventions within a multimodal treatment program. The effect of such a program might differ from the effects of single interventions. Although patients with multimorbidity might be thought less likely to receive quality care based on multiple guidelines, evidence suggests that in practice they might receive better care (23).

Research tends to focus on treatment rather than diagnosis. Most evidence that is used to support recommendations in clinical guidelines concerns therapeutic interventions, and evidence for recommendations about prognosis and diagnosis is often weak (24). Furthermore, single diagnostic tests are usually evaluated, whereas in practice several tests are used.

**No time, no competence**

EBM has been defined as the conscientious, explicit, and judicious use of current best evidence. It aims to integrate evidence with clinical expertise. However, clinicians feel a lack of time and competence to search for, identify, read, assess the quality of, and apply the evidence for each individual patient (25). EBM offers a solution in the form of systematic reviews and clinical guidelines. However, there are now so many available that it may still be impossible for clinicians to keep up with the latest and most valid information. There is evidence from a range of countries and settings that use of EBM by practising clinicians is restricted more by perceived lack of time, skills and knowledge to apply it in practice than by scepticism about the concept itself (e.g. 26,27).

**Little evidence of what works**

In the field of low back pain, EBM rarely provides the basis for action in a primary care consultation. There are plenty of systematic reviews, but they often
conclude that the trials are not good enough or that effect sizes in those of reasonable quality are too small to be of much help. With so little clear and reliable evidence of what works in the first place, it is argued that guidelines often say what you should not do instead of recommending what you should. Thus clinicians may feel that guidelines do not help in their decision-making (21). Studies suggesting “no effect” of a treatment are often too few in number to conclude this is real and consistent.

It is true that for many topics the evidence is sparse or of poor quality (28). But that is no reason for throwing the baby out with the bath water. We have to start somewhere, and have the courage to go on and believe that evidence will slowly accumulate. Identifying gaps in the literature is important and helps to guide priorities for the next step in the often painstaking accumulation of research evidence.

One implication of a commitment among policy makers to basing clinical practice on evidence is that it highlights the continuing scarcity of high quality evidence relative to the volume of practice. Gnanalingham et al (29) reviewed three decades of research publications in prominent journals for five clinical specialities and noted the continuing small proportion which reported on clinical trials. A recent historical review of heart disease recommendations in USA noted that, whilst the number of guideline recommendations for cardiologists had increased, so also had the proportion based on lower levels of evidence or on clinical opinion (28).

A second implication is that guidelines will be skewed to select areas where the evidence has been gathered, and away from interventions for which, however well established, the evidence base is as yet undeveloped.

Small effects only

Another conclusion from back pain research in primary care is that “everything works”. Most people who come into primary care with a new episode, recurrence or exacerbation of low back pain will feel less pain and be more active in a week or two regardless of treatment. RCTs of specific treatments in this situation are trying to find evidence that one thing gives a little more pain relief or promotes a little more activity than another. Keller et al (30), for example, highlighted the range of positive but small effect sizes in trials of non-surgical interventions for back pain. But that fractional extra effect is often only a small proportion of all the improvement everyone is experiencing. This overall improvement reflects the favourable natural course of the condition, explained in part by the effect of seeing and talking with a clinician or a confident optimistic therapist. This is the “placebo effect”, or more precisely the “context effect” (31), highly valued by patients.

This effect is important but it does not remove the need to identify specific advances and improvements that can progress current care. Back pain is very common and additional small beneficial effects in treatment may have substantial population impact, important for both clinical and economic reasons. Although it may be tough to question well-loved treatments on the basis that the same effect could far more easily and cheaply come from helpful and friendly advice, we should still aim as a society to pursue such truth instead of maintaining illusions about our treatments (32).
Lack of generalisability

Uniformity in clinical practice is encouraged on the basis of average results taken from populations included in original studies and systematic reviews. But populations in clinical practice are heterogeneous, and individualised treatment is needed. However research evidence cannot completely fit the profile of an individual patient, although n-of-1 trials might sometimes achieve this. EBM emphasises the need to evaluate the generalisability of research evidence for application to individuals but there will always be uncertainty in doing this. Furthermore RCTs often investigate treatment efficacy in well-controlled experimental circumstances, and effectiveness may be lower in routine circumstances of health care.

The potential for synthesising evidence across different trials has increased. Advanced secondary analysis of existing trials is likely to become as important, perhaps more so, than simply doing more trials. Examples include meta-analysis of individual patient data and Bayesian analysis designed to model outcome at subgroup or individual level. Hayden et al (33,34) conducted a systematic review which concluded that exercise therapy is effective for chronic low back pain. Bayesian analysis additionally identified the optimal type of exercise program and the subgroups most likely to benefit.

As health care databases increasingly provide protected access to well-validated, continuous and complete records of interventions and outcomes in everyday practice, so the potential will grow for RCTs embedded in practice to be combined with observational modelling of routinely collected outcomes of care. This will address current concerns that we are often doing high quality experiments and poor quality generalisations (35).

Other aspects besides evidence are important

Evidence alone is never sufficient to make a clinical decision. Even if the evidence is clear, individuals vary in how they weigh benefits and risks, the final decision influenced by clinicians’ and patients’ values, preferences and expectations, and the inconvenience, availability, and costs of the treatment. Such characteristics now have a different and developing type of research evidence base, including qualitative methodologies which have different objectives and pose different challenges of interpretation for the practising clinician (36).

Different people may make different recommendations, even if they agree on the evidence. UK, US, New Zealand and Dutch national guidelines on low back pain were published almost at the same time (1994 to 1996) and reached similar conclusions regarding the effectiveness of spinal manipulation (37). However, the recommendations varied. Three guidelines recommended spinal manipulation for acute low back pain, but the Dutch guideline did not. The positive guidelines were multidisciplinary, the Dutch guidelines were developed by and for general practitioners only. Weighting of small benefits, potential adverse effects, availability and costs, influenced the Dutch conclusions. This international inconsistency has continued in more recent guidelines (38).

Conflicts of interest

There is always a potential conflict of interest in systematic reviews and clinical guidelines. Sometimes reviews are used to promote a specific intervention. There are more reviews than original RCTs on manipulation for low back pain and neck pain.
Many were conducted by professional groups who may have an interest, explicitly stated or not, conscious or unconscious, in showing that spinal manipulation is effective. Evidence about behavioural treatment for chronic low back pain shows a small difference over waiting list controls but none compared with other interventions (41). This weak finding was used to promote behavioural treatment in some guidelines, but not in others.

Clinical practice guidelines can be seen as a method for cost control by health insurance companies or by the government in countries with a fully nationalised health service such as the UK. Lack of evidence may be used as an argument to exclude an intervention from public health insurance or public funding (42). This can only be justified if there is evidence of no effect. A decade ago commentators expressed concern that evidence-based purchasing of health care might end in disillusion because insufficiency of evidence would deflate expectations of what could be achieved (realised to some extent with the decisions of NICE about cost-effectiveness of new treatments), and that purchasing based on best evidence does not inevitably reduce costs of health care (43).

There are also potential conflicts of interest due to industry or commercial influences (44), and clinicians may feel guidelines are not objective and are oppressive rather than supportive. Recommendations depend on the context in which evidence is being interpreted. Where physiotherapists were plentiful (the Netherlands), evidence-based recommendations reduced the number of physiotherapy sessions. Where there were fewer physiotherapists (the UK), the same evidence was used to argue for increasing their number. In Denmark the same evidence was used to support specific physical therapies; in Germany spas were preferred. Because there is always a potential financial or professional conflict of interest, not only for pharmaceutical companies but also for any health care professional involved in research or guideline development, there is a danger that EBM may be abused for marketing reasons or influenced by subtly selective interests. In the case of evidence that does not support some interventions, more trials may be recommended rather than withdrawal of these treatments.

Two recent initiatives aim at transparent reporting of potential conflicts of interest in systematic reviews and meta-analyses (45) and clinical guidelines (46).

Promoting the industry

Commitment only to things with an evidence base gives a limited view of what is needed in clinical practice. It may promote interventions that are easier to study such as pills; it may ignore interventions for which trial funding is difficult to obtain. Pharmaceutical trials are more numerous and easier to do than non-pharmaceutical trials because of better funding, simpler designs, and strict regulations demanding randomised trials before inclusion in the public health insurance system. Approximately 80% of all trials are industry funded (44).

However the picture is changing, as trials of complex interventions are increasingly used to inform policy and practice, even if the resulting decisions may still be open to influence on the interpretation and marketing of the evidence (47).

Finally there is the evolving issue of litigation related to EBM and guideline use, notably in the US, where they have been both pilloried (48) and supported (49) in court as the basis for litigation against clinicians and their employers.
Hampers innovation

If implementation of new interventions is justified only if sufficient evidence is available, the process may take decades. This issue becomes especially relevant if one RCT is considered insufficient evidence for implementation. The first trial on an intervention is often positive and later trials less positive. The first trial comparing lumbar fusion surgery to conservative care found surgery more effective, but later trials failed to confirm this (50-2).

Conclusions

Criticisms should be taken seriously and addressed. EBM proponents need to better explain that the principles of EBM include room for clinical expertise and other types and sources of information on which to base decision-making in practice. A serious problem, difficult to tackle, is the hampering of innovation if EBM is strictly applied to new interventions. This could be resolved by allowing the use of new, innovative interventions at an early stage within the setting of randomised controlled trials, or observational, monitoring or audit studies. The alternative is to define characteristics of interventions or situations that would be acceptable as exceptions to EBM.

Clinicians should use guidelines only in situations where they do not immediately know what clinical decision to make. In routine cases, guidelines are not needed. The number and size of systematic reviews and clinical guidelines is a growing and serious problem. The solution probably lies in effective and efficient implementation strategies, which do not reduce the patient-clinician relationship to a dull and potentially negative uniformity. Developing these strategies is a major challenge for the future.

EBM users and followers often focus only and strictly on the evidence, and even more specifically on evidence from RCTs. However, ‘true’ EBM acknowledges that evidence is important, but that other aspects play a role as well in making final clinical or policy decisions - patients’ and clinicians’ preferences, availability of treatments, or financial, ethical and legal issues. A major challenge to current EBM is how to evaluate and integrate different types of research from a range of conceptual and methodological backgrounds into frontline clinical care and shared decision-making with patients. This is often reflected in guideline development, where these other aspects can be taken into account.

There are useful arguments both for and against EBM. In our opinion, the problem is not so much that the principles of EBM are flawed, but applying EBM in clinical or policy practice is challenging. It is better to make well-informed decisions, either as a clinician making decisions in daily practice together with patients, or as a policy maker making decisions that may have a huge health and financial impact on a large population.

We hope that the discussion of the pros and cons of EBM will continue in the spirit of weighing and considering, rather than contradicting or taking for granted.

References

35. Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?” Lancet 2005;365:82-93