Who will challenge evidence-based medicine?

N W Goodman

ABSTRACT — Evidence based medicine (EBM), which lays special emphasis on the formal gathering and synthesis of research data by systematic review and meta-analysis, is intuitively attractive as a way of determining best treatments. However, questions remain unanswered about the underlying assumptions of EBM. There is little true dialogue between its critics and its advocates, although the latter have had considerable disagreements among themselves in the correspondence columns of medical journals about the precise way in which its methods should be used. EBM medicine is in danger of becoming a new and unchallengeable orthodoxy following its own political agenda.

Medical practice is best based on evidence — without special emphasis that any particular form of evidence is in all senses better — but the core of evidence-based medicine (EBM) is the underlying process of randomised controlled trials (RCTs), systematic review and meta-analysis. This gives EBM its special meaning and, to some, a special authority. A small industry has arisen from it, and there are many whose livelihoods now depend upon EBM.

The specialties of neonatology and obstetrics have been an important springboard for EBM but, in discussion with some trainees, I was astounded to learn that, while they knew of systematic reviews that had not been well received, they did not realise that the assumptions of EBM could be anything other than valid. One of the reasons for EBM was the presumed opinionated dogma of the expert. Has this dogma been replaced by dogma from a different source?

Lack of self-doubt

A new medical treatment requires evidence of effectiveness from prospective studies. Surely this is even more true when EBM is put forward as a new, and in all ways better, method of deciding best treatment? Yet this evidence is lacking; it is taken as self-evident, but it is not. Charlton wrote:

The basic error of EBM is quite simple. It is that epidemiological data do not provide the information necessary to treat individual patients. The error is intractable and intrinsic to the methodological nature of epidemiology, and no amount of statistical jiggery-pokery with huge data sets can make any difference.

Charlton is not the only critic of the assumptions of EBM; Feinstein’s criticisms are especially scholarly. To give analogy, it is as if a fashion house collected data from many samples of women, concluded that the average size was 16 (with large numbers conferring small confidence limits), and thus restricted production to clothes only of size 16.

Sackett and co-authors’ book is a popular manual of EBM. Neither Feinstein nor any other critic is cited in this book; it contains no justification, only presumption that following the schemes outlined therein will be better for patients. This is not an isolated occurrence. Hampton was unhappy that a book on evidence-based cardiology is strong on published trials but weak on philosophy. The reference lists to a series of articles in the British Medical Journal ‘analysing the gap between research and practice’ include many publications of its authors and of other names well known in EBM circles, but no critical publications. Even the article entitled Barriers and bridges to evidence based clinical practice fails to acknowledge the possible barrier of invalidity, although it is admitted (in a different article) that:

In many cases the implications of research evidence for policy and practice are not straightforward or obvious, – a statement made with citation to an article by Naylor about ‘grey zones’ of medicine.

The series of articles in the British Medical Journal was adapted from a book published simultaneously by the BMJ publishing group. EBM has powerful friends: duplicate publication is frowned upon, and editors reject material published elsewhere – but not if the message is deemed sufficiently ‘important’. With EBM accepted (indeed, it has been suggested that revalidation may depend on consultants practising it), there is an almost infinite number of possible publications, each further justifying the acceptance. The critics, meanwhile, are largely restricted to repeating themselves, mainly in the correspondence columns – where their opinions can be safely ignored.

However, if the critics are correct that ‘the error is intractable and intrinsic’, then no matter how many people are producing no matter how large a body of work, it is a mirage.

Not all the advocates of EBM are so blinkered. Davey Smith’s group recognised difficulties and dissensions, but did not answer the basic criticism. At present, I do not believe that the advocates have satisfactorily answered their critics. As an example, Shapiro, writing of putative epidemiological hazards such as living near nuclear installations, proposed that ‘meta-analysis of published nonexperimental data should be abandoned’, warning that:

Despite the recognition that the possibilities of bias can seldom be eliminated and that possible confounding [factors]
never can be [eliminated], some meta-analysts have proceeded on the assumption that exactly those objectives can be achieved by assigning appropriate quality weights...\(^{11}\)

Egger et al write ‘We disagree’, but:

the statistical combination of studies should not generally be a prominent component of reviews of observational studies.\(^{12}\)

This is not a robust disproof of Shapiro; it is a difference of opinion. Why should we unquestioningly believe either view?

**Technical disagreements about meta-analysis**

Let us assume that EBM is valid. At its heart is meta-analysis, a statistical method, but meta-analysis has some problems.

Davey Smith’s group is keen on a particular graphic representation in meta-analysis known as a funnel plot\(^{13}\), which is one way of making sense of studies of different sizes. Because of its supposed ability to detect publication bias (that is, the failure to report negative trials), Egger et al\(^{14}\) suggest a statistic derived from funnel plot analysis be integral to meta-analysis. In the correspondence, there was much discussion of publication bias, of the importance of being systematic, and of meta-analysis not being a panacea, all of which gave the curious impression that some people are truly surprised that medical knowledge is so recalcitrant. A pair of papers in the *Lancet*\(^{15,16}\) provoked similar unsatisfactory disagreements in the ensuing correspondence.

If meta-analysis really is so difficult, are we right to trust it? Perhaps these and other issues explain the confusion of Prins and Büller\(^{17}\). Four meta-analyses of aminoglycoside dosage schedules reached different conclusions about efficacy and side effects. These authors commented that all physicians could do was to choose the conclusion to fit their own belief.

**Beyond meta-analysis**

A key issue is whether meta-analyses predict later valid findings\(^{18}\). EBM is in trouble if even some (how many?) meta-analyses fail to predict later valid findings (though how would we know they were valid?). When LeLorier et al\(^{19}\) showed that meta-analysis failed to predict the finding in a subsequent trial in one-third of 40 compared outcomes, Bailar commented in an accompanying editorial that LeLorier’s study also cast doubt on the validity of large trials: is the ‘truth’ represented by the trial or by the meta-analysis? Bailar wrote\(^{20}\):

> I know of no instance in medicine in which a meta-analysis led to a major change in policy before the time when a careful, conventional review of the literature led to the same change.

LeLorier’s paper set the correspondence columns alight\(^{21}\). Most letters were critical, pointing out methodological flaws in LeLorier’s approach, or accusing Bailar of only voicing an opinion. LeLorier and colleagues replied in detail, but Bailar remained unrepentant. He admitted that his objections to meta-analysis were ‘purely pragmatic’, justifying this on the grounds that meta-analysis:

> does not work nearly as well as we might want it to work. The problems are so deep and so numerous that the results are simply not reliable.

While not denying that people want to do things well, Bailar saw the insurmountable problem that ‘meta-analysis often fails in ways that seem to be invisible to the analyst’. He repeated his belief in the:

> knowledgeable, thoughtful, traditional review ... the closest thing we have to a gold standard of summarizing disparate evidence in medicine.

A conversation between Bailar and Davey Smith would be interesting, but I fear it might not be enlightening

> ... [meta-analysis] is clearly superior to the narrative approach to reviewing medical research\(^{22}\).

The paper of LeLorier et al generated heat in other places when other journals made editorial comment\(^{23,24}\). Davey Smith and Egger repeated their description of funnel plots\(^{25}\). LeLorier and Grégoire disagreed with ‘several’ of Naylor’s comments\(^{26}\). Naylor and Davey Smith in turn\(^{27}\) reckoned that LeLorier and Grégoire had ‘missed the salient methodological issues’, and ended their letter by suggesting that any later comparison is invalid because the whole point of meta-analysis is to consider the totality of the evidence. If later trials are distinct, they should neither be compared nor combined. This seems the ultimate rejection by the EBM advocates of any test of EBM: new evidence is valid only if added to existing evidence.

**Conclusion**

There is an *ideological difference of opinion* about EBM. As there is no evidence (and unlikely ever to be) that EBM provides *better medical care in total* than whatever we like
to call what went before, the debate will continue, as most human debates do, on emotional and rhetorical grounds.

Writers whose arguments seem to me to be valid have doubts about EBM. These doubts may be simply about over-rigorous application of the obtained numbers. For example, Kleinert wrote28 that:

the modern trend to search for precise answers in the form of numbers and probabilities can have only a limited role in human sciences such as medicine.

The doubts extend to unease at the way some people, including politicians, equate EBM to the whole of medicine29, and to complete dismissal of the whole process30. For some conditions, there is no doubt that what happens in RCTs is not what happens in the real world4. This must therefore be true for at least some meta-analyses, because they are based largely on RCTs. Swales31 noted that the 'proven' advantage of surgery as treatment for carotid artery stenosis was entirely negated by the 9.8% complication rate in community practice: the complication rate in the studies was 3.7%. In the wider sense, EBM must fail because it cannot take account of societal and moral decisions28.

It is all very well writing that:

The technological barriers to worldwide data exchange and collaboration are tumbling down – we can only hope that the remaining barriers, rooted in customary practice, political agendas, and commercial interests, will swiftly fall too32.

These are worthy thoughts, but first we need to be sure that EBM is not simply another 'customary practice', following its own political agenda, and generating its own commercial interests.

References

1. Charlton BG. Book review of evidence-based medicine. J Eval Clin Pract 1997;3:169–72.
2. Feinstein AR. Clinical judgement revisited: the distraction of quantitative models. Ann Intern Med 1994;120:799–805.
3. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence based medicine. How to practice and teach EBM. New York: Churchill Livingstone. 1997.
4. Hampton JR. Evidence based cardiology. Br Med J 1998;317:132–6.
5. Haynes B, Haines A. Barriers and bridges to evidence based clinical practice. Br Med J 1998;317:273–6.
6. Sheldon TA, Guyatt GH, Haines A. When to act on the evidence. Br Med J 1998;317:139–42.
7. Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. Lancet 1995;345:840–2.
8. Haines A, Donald A (eds). Getting research findings into practice. London: British Medical Journal Publishing Group, 1998.
9. Bhopal RS, Tomàs A. The role of letters in reviewing research. Br Med J 1994;308:15823.
10. Egger M, Davey Smith G. Meta-analysis: bias in location and selection of studies. Br Med J 1998;316:61–6.
11. Shapiro S. Meta-analysis/Shmata-analysis. Am J Epidemiol 1994;140:7718.
12. Egger M, Schneider M, Davey Smith G. Meta-analysis: spurious precision? Meta-analysis of observational studies. Br Med J 1998;316:140–4.
13. Light RJ, Pillemer DB. Summing up: the science of reviewing research. Cambridge, MA: Harvard University Press, 1984.
14. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Br Med J 1997;315:629–34.
15. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. Lancet 1998;351:47–52.
16. Lau J, Ioannidis JPA, Schmid CH. Summing up evidence: one answer is not always enough. Lancet 1998;351:123–27.
17. Prins JM, Bùller HR. Meta-analysis: the final answer, or even more confusion? (letter) Lancet 1996;348:199.
18. Chalmers I, Stewart L.A. Can meta-analyses be trusted? (letter) Lancet 1991;338:1464–5.
19. Le Lorier J, Grégoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med 1997;337:536–42.
20. Bailar III JC. The promise and problems of meta-analysis. N Engl J Med 1997;337:559–61.
21. Correspondence. Meta-analyses and large randomized, controlled trials. N Engl J Med 1998;338:59–62.
22. Davey Smith G, Egger M. Unresolved issues and future developments. Br Med J 1998;316:221–5.
23. Anon. Meta-analysis under scrutiny. Lancet 1997;350:675.
24. Naylor CD. Meta-analysis and the meta-epidemiology of clinical research. Br Med J 1997;315:617–19.
25. Davey Smith G, Egger M. Meta-analysis of randomised controlled trials (letter). Lancet 1997;350:1182.
26. Le Lorier J, Grégoire G. Meta-analysis and the meta-epidemiology of clinical research. Comments on paper by author of editorial were unwarranted (letter). Br Med J 1998;316:311–2.
27. Naylor CD, Davey Smith G. Test meta-analyses for stability. Br Med J 1998;317:206–7.
28. Kleinert S. Rationing of health care – how should it be done? Lancet 1998;352:1244.
29. Black D. The limitations of evidence. J R Coll Physicians Lond 1998;32:233–4.
30. Charlton BG, Miles A. The rise and fall of EBM. Q J Med 1998;91:371–4.
31. Swales J. The National Health Service and the science of evaluation: two anniversaries. Health Trends 1998;30:20–2.

Address for correspondence: Dr N W Goodman, Department of Anaesthesia, Southmead Hospital, Bristol BS10 5NB