

This is a repository copy of *Getting research findings into practice : when to act on the evidence.*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/956/>

Article:

Sheldon, T A orcid.org/0000-0002-7479-5913, Guyatt, G H and Haines, A (1998) Getting research findings into practice : when to act on the evidence. *BMJ*. pp. 139-142. ISSN 1756-1833

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

*Getting research findings into practice***When to act on the evidence**

Trevor A Sheldon, Gordon H Guyatt, Andrew Haines

There is increasing interest in providing evidence based health care—that is, care in which healthcare professionals, provider managers, those who commission health care, the public, and policymakers consistently consider research evidence when making decisions.^{1,2} Purchasers, for example, should be able to influence the organisation and delivery of care (such as for cancer³ and stroke services⁴) and the type and content of services (such as using chiropractic for back pain or dilatation and curettage and drug treatment for menorrhagia⁵). Policymakers should ensure that policies on treatment reflect and are consistent with research evidence, and that the incentive structure within the health system promotes cost effective practice. They must also ensure that there is an adequate infrastructure for monitoring changes in practice and for producing, gathering, summarising, and disseminating evidence. Clinicians determine the day to day care patients receive in healthcare systems, and user groups (for example, patients, their families, and their representatives) are also beginning to play an important role in influencing healthcare decisions.⁶

The factors described below should be considered when deciding whether to act on or promote the implementation of research findings.

Convincing evidence of net benefit**Evaluating the methods of primary studies**

Individual research studies vary in their degree of bias—that is, how much they are likely to underestimate or overestimate the effectiveness of an intervention. Observational studies, in which investigators compare the results of groups of patients who are receiving different treatments based on the patient's own or the clinician's preference, are susceptible to bias because the prognosis of the groups is likely to differ in unpredictable ways, leading to spuriously reduced or, more commonly, inflated treatment effects.

Rigorous randomised control trials greatly reduce bias by ensuring that the groups being compared are similar.⁷ As long as patients are analysed in the groups to which they were randomised, this type of trial permits a more confident inference that the treatments offered are responsible for differences in outcome. Randomised controlled trials are useful not only for testing the effectiveness of interventions in tightly controlled clinical settings but also across a wide spectrum of health research.^{8,9} Inferences are further strengthened if patients, care givers, and those assessing outcomes are blind to the allocation of patients to treatment or control groups and if follow up is complete.¹⁰

While randomised controlled trials are often regarded as the gold standard for comparing the efficacy of treatments, other study designs are appropriate for evaluating other types of healthcare technologies, such as diagnostic tests, or for assessing

Summary points

There is increasing interest in making clinical and policy decisions based on research findings

Not all research findings should or can be implemented; prioritisation is necessary

The decision whether to implement research evidence depends on the quality of the research, the degree of uncertainty of the findings, relevance to the clinical setting, whether the benefits to the patient outweigh any adverse effects, and whether the overall benefits justify the costs when competing priorities and available resources are taken into account

Systematic reviews that show consistent results are likely to provide more reliable research evidence than non-systematic reviews or single studies

Researchers should design studies that take into account how and by whom the results will be used and the need to convince decision makers to use the intervention studied

the potentially harmful effects of interventions.¹¹ Qualitative methods are increasingly being used, for example, to provide an understanding of patients' and professionals' attitudes and behaviours, the effects of culture, the context of healthcare, and their interactions.¹²

Whatever the appropriate design, practitioners will often discover that research evidence is biased or otherwise limited; for example, the investigators may have focused on inappropriate physiological end points rather than outcomes relevant to patients.¹³ In evaluations of the organisation of health care, providers must consider whether treatment effects were really due to the putative intervention; for example, in randomised controlled trials that found a positive effect of stroke units, was the impact really due to the organisational structure or to the greater skill or enthusiasm of those who established the units?¹⁴ Though practitioners will still need to use imperfect research information, new clinical policies should not be implemented unless clinicians find that there is strong evidence of benefit.

Evaluating the methods and results of systematic reviews

Systematic reviews can provide reliable summaries of data that address targeted clinical questions; they can also provide less biased estimates of treatment effects if they adhere to the criteria shown in the box.¹⁴

This is the second in a series of eight articles analysing the gap between research and practice

NHS Centre for Reviews and Dissemination, University of York, York YO1 5DD
Trevor A Sheldon,
professor

Departments of Medicine, and Clinical Epidemiology and Biostatistics, McMaster University, 1200 Main Street West, Hamilton, Ontario, Canada L8N 3Z5
Gordon H Guyatt,
professor

Department of Primary Care and Population Sciences, Royal Free and University College London Schools of Medicine, London NW3 2PF

Andrew Haines,
professor of primary health care

Correspondence to: Professor Sheldon
tas5@york.ac.uk

Series editors: Andrew Haines and Anna Donald

BMJ 1998;317:139–42

Criteria that increase the reliability of a systematic review

- Use of explicit criteria for inclusion and exclusion; these should specify the population, the intervention, the outcome, and the methodological criteria for the studies included in the review
- Use of comprehensive search methods to locate relevant studies, including searching a wide range of computerised databases using a mixture of appropriate key words and free text
- Assessment of the validity of the primary studies; this should be reproducible and attempt to avoid bias
- Exploration of variation between the findings of the studies
- Appropriate synthesis and, when suitable, pooling of primary studies

A rigorous systematic review may sometimes leave the decision maker who is reading it uncertain. Classification of the strength of research evidence should consider each of the following four points. Firstly, the methodology of the primary studies may be weak. Secondly, unexplained variability between study results may lead to doubt about the results of studies that show larger treatment effects or those that show no benefit. Thirdly, small sample sizes may lead to wide confidence intervals even after results have been pooled across studies. Thus, the research evidence may be consistent with a large or a negligible treatment effect. Fourthly, because of the side effects associated with a treatment, or their cost, the balance between treating and not treating with an effective intervention may be precarious.

Grades of the strength of the evidence of the effectiveness of a treatment have been developed that account for the type and quality of the study design and the variability of study results.¹⁵ Thus, a systematic review of randomised controlled trials that show consistent results (such as trials of streptokinase for treatment of acute myocardial infarction²) would be graded as providing higher quality evidence than a review of randomised controlled trials that show variable results without a good explanation of the variability (heterogeneity).

The precision of the estimated treatment effect, and the trade off between the benefits and risks could also

be considered. When assessing risks it is important to note that many studies of efficacy, and reviews of these studies, do not provide sufficient information about the possible harm of treatments. Sample sizes in most randomised trials are usually not large enough and the study period not long enough to detect rare or long term harmful effects.¹⁶ Large observational studies may be useful in determining the probability of harm.¹⁷

Putting evidence of benefit into perspective

Evidence of effectiveness alone does not imply that an intervention should be adopted; adoption of an intervention depends on whether the benefit is sufficiently large relative to the risks and costs. For example, the small positive effect of interferon beta in the treatment of multiple sclerosis relative to its cost makes implementation of its use questionable.¹⁸

One approach to the decision about whether an intervention should be implemented is to determine a threshold above which treatment would routinely be offered and below which it would not. Decision makers might consider the threshold in terms of the number of patients one would need to treat to prevent a single adverse event (such as a death).¹⁹ The threshold number needed to treat defines the value above which the disadvantages of treatment outweigh the benefits (and treatment may therefore be withheld), and below which the benefits outweigh the disadvantages (and treatment may therefore be offered).²⁰ Because the cost of treatment and the benefit to the length and quality of life vary, each intervention needs a separate threshold; this threshold will also vary according to the values of the patient, or population, being offered the intervention.

When reliable data are available, a threshold might be expressed in terms of a cost effectiveness ratio that defines the cost of achieving a unit of benefit below which an intervention is seen as worth implementing routinely (for example, quality adjusted life years that take social values about the equity of health and resource allocation into account). Quantitative research evidence is inevitably probabilistic and subject to various forms of uncertainty; it is rarely the sole basis of decision making at the governmental or clinical level. Indeed, uncertainty is one obstacle to policymakers using research evidence.²¹ People differ in their willingness to take risks; these differences explain the variations in decisions made when the same evidence is evaluated by different people. However, research evidence should play an important, and greater, part in decision making and can provide a benchmark against which decisions can be audited.

Applying research to practice

Whether research evidence can or should be applied to a specific patient cannot always be deduced straightforwardly from the research. Results of evaluative studies are usually given as average effects. Patients may differ from the average in ways that influence the effectiveness of the treatment (relative risk reduction) or its impact (absolute risk reduction).^{22 23} Factors that clinicians and patients should consider before applying research evidence to a specific case are summarised in the box.



Factors to consider when applying evidence to individual patients

- Is the relative risk reduction that is attributed to the intervention likely to be different in this case because of the patient's physiological or clinical characteristics?
- What is the patient's absolute risk of an adverse event without the intervention?
- Is there significant comorbidity or a contraindication that might reduce the benefit?
- Are there social or cultural factors that might affect the suitability of treatment or its acceptability?
- What do the patient and the patient's family want?

Patients who participate in trials may not be typical of the types of the people for whom the treatment is potentially useful.²⁴ None the less, it is probably more appropriate to assume that research findings are generalisable across patients unless there is strong theoretical or empirical evidence to suggest that a particular group of patients will respond differently.²²

There may be a heterogeneity of effect across patients because of biological, social, or other differences that influence the effect of the intervention or the risk of an adverse outcome.^{24 25} For example, β blockers may be less effective than diuretics in lowering blood pressure in black people of African descent than in white populations.²⁶ Interventions are more likely to have a uniform impact when the effect of treatment is purely a biological process, and where there is less variation within the population than when many factors specific to the patient or specific to the context mediate the effect.²⁷ The issue of whether treatment effects are constant or are likely to be sensitive to patient and context is important when targeting effective treatments to economically disadvantaged groups of people with the aim of reducing inequalities in health. If, for example, smoking cessation interventions are less successful in poorer people, then such programmes might not have the anticipated effects on health equity.

Single patient randomised controlled trials (n of 1 trials) may help determine a particular patient's response to treatment in a number of chronic conditions, including chronic pain syndromes such as arthritis or chronic heart or lung disease, in which the benefit of treatment may vary widely between individual patients.²⁸

Clinicians must carefully consider treatments in patients for whom treatment may be contraindicated or where there is substantial comorbidity. In patients with comorbid conditions, a reduction in the risk of dying from one disease might not reduce the overall risk of dying because of the risk of a competing cause of death.

The effect of an intervention may also vary because patients do not share the same morbidity or risk.²⁹ For any given measurement of the effectiveness of treatment patients at higher risk will generally experience greater levels of absolute risk reduction or impact from treatment.^{25 29-31} For example, patients at high risk of dying from coronary heart disease who are treated with drugs to lower cholesterol will experience a greater reduction in the risk of dying than those at lower risk—that is, 30 patients at high risk might have

to be treated for five years to save one life, but 300 patients at low risk would have to be treated to save one life.^{32 33} Thus, a treatment that might be worth implementing in a patient at high risk may not be worth implementing in a patient at lower risk.^{32 33}

The decision whether to use a treatment also depends on factors that are specific to the patient. Clinicians will find that research studies that consider a range of important outcomes of treatment are more useful than those which have only measured a few narrow clinical end points. More qualitative research done within robustly designed quantitative studies will help practitioners and patients to better understand and apply the results of research.

Setting priorities

Implementation of research evidence occurs rarely unless there are concerted attempts to get the results into practice.³⁴ It is impossible to promote actively the implementation of the results of all systematic reviews because of the limited capacity of healthcare systems to absorb new research and the investment necessary to overcome the obstacles of getting research into practice. These costs must be considered in relation to the likely return in terms of improvements in health. The anticipated benefits of implementation vary according to factors such as the divergence between research evidence and current practice or the pressure of policies that influence the marginal benefit of further efforts at implementation.

When evaluating the same evidence different decision makers will use different criteria to prioritise treatments for implementation. Policymakers, for example, may look for societal gains in health and efficiency, while clinicians may consider the wellbeing of their patients to be most important.³⁵ Formal decision analysis may be helpful in setting priorities for implementation and in applying research evidence to the treatment of individual patients.^{36 37}

The degree to which clinicians see even good quality research as able to be implemented will depend on the extent to which the results conflict with professional experience and beliefs. This reflects an epistemological mismatch between the sort of evidence that researchers produce and believe in and the sort of evidence that practising clinicians value.³⁸ In many cases the implications of research evidence for policy and practice are not straightforward or obvious,³⁹ and this ambiguity may result in the same evidence giving rise to divergent conclusions and actions.⁴⁰ Depending on the perceived risks, the extent of change required, and the quality and certainty of the research results, many clinicians and policymakers will wait for confirmatory evidence. When designing studies investigators should consider how and by whom their results will be used. The design should be sufficiently robust, the setting sufficiently similar to that in which the results are likely to be implemented, the outcomes should be relevant, and the study size large enough for the results to convince decision makers of their importance.

Funding: None.

Conflict of interest: None.

- 1 Davidoff F, Haynes RB, Sackett D, Smith R. Evidence based medicine. *BMJ* 1995;310:1122-6.
- 2 Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of the results of controlled trials and recommendations of clinical experts: treatments for myocardial infarction. *JAMA* 1992;268:240-8.
- 3 Department of Health. *NHS Executive guidance for purchasers: improving outcomes in breast cancer*. London: DoH, 1996. (96CC0021.)
- 4 Stroke Unit Trialists' Collaboration. A systematic review of specialist multidisciplinary (stroke unit) care for stroke inpatients. In: Warlow C, Van Gijn J, Sandercock P, eds. *Stroke module of the Cochrane database of systematic reviews*. London: BMJ Publishing, 1997.
- 5 NHS Centre for Reviews and Dissemination. *The management of menorrhagia*. York: University of York, 1995. (Effective health care bulletin no 9.)
- 6 Entwistle V, Sheldon TA, Sowden A, Watt I. Evidence informed patient choice: issues of involving patients in decisions about health care technologies. *Int J Technol Assess Health Care* 1998 (in press).
- 7 Schultz K, Chalmers I, Haynes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled clinical trials. *JAMA* 1995;273:408-12.
- 8 Green SB. The eating patterns study—the importance of practical randomized trials in communities. *Am J Public Health* 1997;87:541-3.
- 9 Bauman KE. The effectiveness of family planning programs evaluated with true experimental designs. *Am J Public Health* 1997;87:666-9.
- 10 Guyatt GH, Sackett DL, Cook DJ for the Evidence-based Medicine Working Group. Users' guides to the medical literature. II. How to use an article about therapy or prevention. Part A. Are the results valid? *JAMA* 1993;270:2598-601.
- 11 Laupacis A, Wells G, Richardson WS, Tugwell P for the Evidence-based Medicine Working Group. Users' guides to the medical literature. V. How to use an article about prognosis. *JAMA* 1994;272:234-7.
- 12 Mays N, Pope C. *Qualitative research in health care*. London: BMJ Publishing, 1996.
- 13 Gotsche PC, Liberati A, Torri V, Rossetti L. Beware of surrogate endpoints. *Int J Technol Assess Health Care* 1996;12:238-46.
- 14 NHS Centre for Reviews and Dissemination. *Undertaking systematic reviews of research on effectiveness: CRD guidelines for those carrying out or commissioning reviews*. York: University of York, 1996. (CRD report no 4.)
- 15 Liddle J, Williamson M, Irwig L. *Method for evaluating research and guidelines evidence*. Sydney: NSW Health Department, 1996.
- 16 Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V for the Evidence-based Medicine Working Group. Users' guides to the medical literature. IV. How to use an article about harm. *JAMA* 1994;271:1615-9.
- 17 Palareti G, Leali N, Coccher S, Poggi M, Manotti C, D'Angelo A, et al. Bleeding complications of oral anticoagulant treatment: an inception-cohort, prospective collaborative study (ISCOAT). *Lancet* 1996;348:423-8.
- 18 Walley T, Barton S. A purchaser perspective of managing new drugs: interferon beta as a case study. *BMJ* 1995;311:796-9.
- 19 Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment. *BMJ* 1995;310:452-4.
- 20 Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. *JAMA* 1995;274:1800-4.
- 21 Hammond K, Mumpower J, Dennis RL, Fitch S, Crumpacker W. Fundamental obstacles to the use of scientific information in public policy making. *Technological Forecasting and Social Change* 1983;24:287-97.
- 22 Davis CE. Generalizing from clinical trials. *Control Clin Trials* 1994;15:11-4.
- 23 Wittes RE. Problems in the medical interpretation of overviews. *Stat Med* 1987;6:269-76.
- 24 Pearson TA, Myers M. Treatment of hypercholesterolaemia in women: equality, effectiveness and extrapolation of evidence. *JAMA* 1997;277:1320-1.
- 25 Bailey KR. Generalizing the results of randomized clinical trials. *Control Clin Trials* 1994;15:15-23.
- 26 Veterans Administration Co-operative Study Group. Comparison of propranolol and hydrochlorothiazide for the initial treatment of hypertension. I. Results of short-term titration with emphasis on racial differences in response. *JAMA* 1982;248:1996-2003.
- 27 Cowan CD, Wittes J. Intercept studies, clinical trials, and cluster experiments: to whom can we extrapolate? *Control Clin Trials* 1994;15:24-9.
- 28 Mahon J, Laupacis A, Donner A, Wood T. Randomised study of n of 1 trials versus standard practice. *BMJ* 1996;312:1069-74.
- 29 Davey Smith G, Eggar M. Who benefits from medical interventions? *BMJ* 1993;308:72-4.
- 30 Glasziou PP, Irwig LM. An evidence based approach to individualising treatment. *BMJ* 1995;311:356-9.
- 31 Oxman AD, Guyatt GH. A consumer's guide to subgroup analysis. *Ann Intern Med* 1992;116:78-84.
- 32 Davey Smith G, Song F, Sheldon TA. Cholesterol lowering and mortality: the importance of considering initial level of risk. *BMJ* 1993;306:1367-73.
- 33 Marchioli R, Marfisi RM, Carinci F, Tognoni G. Meta-analysis, clinical trials, and transferability of research results into practice. *Arch Intern Med* 1996;156:1158-72.
- 34 Oxman A, Davis D, Haynes RB, Thomson MA. No magic bullets: a systematic review of 102 trials of interventions to help health professionals deliver services more effectively and efficiently. *Can Med Assoc J* 1995;153:1423-43.
- 35 Diamond GA, Denton TA. Alternative perspectives on the biased foundation of medical technology assessment. *Ann Intern Med* 1993;118:455-64.
- 36 Thornton JG, Lifford RJ. Decision analysis for managers. *BMJ* 1995;310:791-4.
- 37 Dowie J. "Evidence-based," "cost-effective" and "preference-driven" medicine: decision analysis based medical decision making is the pre-requisite. *J Health Services Res Policy* 1996;1:104-113.
- 38 Tanenbaum SJ. Knowing and acting in medical practice: the epistemological politics of outcomes research. *J Health Polit Policy Law* 1994;19:27-44.
- 39 Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet* 1995;345:840-2.
- 40 Tanenbaum SJ. "Medical effectiveness" in Canadian and US health policy: the comparative politics of inferential ambiguity. *Health Serv Res* 1996;31:517-32.

The articles in this series are adapted from *Getting research findings into practice*, edited by Andrew Haines and Anna Donald, which will be published in July.

Doctors who changed a patient's life Today's dogma may be tomorrow's joke

She was a tall, distinguished looking, elderly woman who, despite having severe nephrotic syndrome, still carried herself erect. My inquiry about her medical history elicited a startling response.

Fifty years previously she had been an attractive, athletic young woman, engaged to be married, when she developed severe peripheral oedema. Her wealthy parents spared no expense in obtaining the best medical opinions available. All the distinguished doctors agreed that she had nephritis, that she should have strict bed rest in a darkened room, and that the prognosis was grave, telling her parents that she had, at most, six weeks to live.

After several weeks of full nursing care, it was apparent that not only was she not dead, but that she was also no longer oedematous. The experts again gathered and opined that she had indeed been fortunate, but that she would be delicate for the rest of her short life, and should avoid any physical exertion. They particularly emphasised that pregnancy would be extremely dangerous.

The marriage went ahead, but the plans were drastically altered. The couple bought a bungalow with a small garden rather than the rather grand property that they had planned, so that she need not exert herself. My patient and her husband realised that abstinence was the only effective contraception, so their long marriage was apparently never consummated.

Despite living the life of a semi-invalid, she did not have a further day's illness for 30 years, when she again developed severe oedema. She consulted a famous doctor, whose specialty was not

nephrology. Without much of an investigation, he prescribed steroids with dramatic effect. Twenty years later, for her third episode of the nephrotic syndrome, I performed a renal biopsy and countless other tests before prescribing steroids. The biopsy showed minimal change nephropathy, a condition characterised by relapses and remissions, either spontaneously or with the help of steroids.

As doctors, we sometimes forget that our utterances may be explicitly believed. The more dogmatic the statement and the more distinguished the doctor, the more likely it is that instructions may be slavishly obeyed. While the advice given to my patient may have been correct by the knowledge of the day, it was clearly nonsense, and it ruined her life. We should remember that today's dogma may be tomorrow's joke, and that its debunking may inadvertently leave some casualties in its wake.

John H Turney, consultant physician and nephrologist, Leeds

We welcome articles up to 600 words on topics such as *A memorable patient, A paper that changed my practice, My most unfortunate mistake*, or any other piece conveying instruction, pathos, or humour. If possible the article should be supplied on a disk. Permission is needed from the patient or a relative if an identifiable patient is referred to.