This is an author produced version of *Getting research findings into practice*.

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

**Article:**
Sheldon, T A, Guyatt, G H and Haines, A (1998) Getting research findings into practice. BMJ. pp. 139-142. ISSN 1756-1833
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. Purchasers, for example, should be able to influence the organisation and delivery of care (such as for cancer and stroke services) and the type and content of services (such as using chiropractic for back pain or dilatation and curettage and drug treatment for menorrhagia). 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.

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. 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.

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.

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 the potentially harmful effects of interventions. 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.

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. 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? 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.

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

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 Y01 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 tash@york.ac.uk
Series editors: Andrew Haines and Anna Donald

BMJ 1998;317:139–42
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 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. Thus, a systematic review of randomised controlled trials that show consistent results (such as trials of streptokinase for treatment of acute myocardial infarction) 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. Large observational studies may be useful in determining the probability of harm.

**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 costs and harms. 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.

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). 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). 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. 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). Factors that clinicians and patients should consider before applying research evidence to a specific case are summarised in the box.

---

**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.
Patients who participate in trials may not be typical of the types of the people for whom the treatment is potentially useful. 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.

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. For example, β blockers may be less effective than diuretics in lowering blood pressure in black people of African descent than in white populations. 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. 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.

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. 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. 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. 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.

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. 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. Formal decision analysis may be helpful in setting priorities for implementation and in applying research evidence to the treatment of individual patients.

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. In many cases the implications of research evidence for policy and practice are not straightforward or obvious, and this ambiguity may result in the same evidence giving rise to divergent conclusions and actions. 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.
Doctors who may be tomorrow’s joke

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 stammering response. Fifty years previously she had been an attractive, athletic young woman, engaged to be married, when she 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 heard. 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