Applying the results of clinical trials to patients in general practice: perceived problems, strengths, assumptions, and challenges for the future

TOM FAHEY

SUMMARY

Randomized controlled trials (RCTs) and systematic reviews of RCTs now provide the most robust external evidence about the effectiveness of patient care. There are, however, several assumptions made when applying the results of RCTs to individual patients. This paper aims to outline the perceived barriers against the use of RCTs in practice, while emphasizing the rationale and advantages underlying the approach. A critical discussion concerning the assumptions made when applying evidence from RCTs to individual patients will be presented, with a worked example derived from a patient with acute sinusitis. Finally, proposals concerning the effective implementation of evidence derived from RCTs in the context of individual patient care will be discussed.

Keywords: randomized controlled trials; patient care; acute sinusitis; evidence-based medicine.

Introduction

THE realization that epidemiology has a direct part to play in individual patient care has led to the emergence of clinical epidemiology, now known as evidence-based medicine (EBM).[1-4] Randomized controlled trials (RCTs) are now regarded as the most reliable method by which a treatment can be assessed.[5] Systematic reviews of RCTs and their quantitative pooling by means of meta-analysis have evolved more recently and provide further evidence concerning the effectiveness of treatments.[6,7] But where does all this progress and information concerning the value of different treatments leave a general practitioner (GP) when faced with an individual patient's clinical problem? A common perception is that extrapolating the results of an RCT to an individual patient requires 'a great leap of faith'.[8] Is this view justified? By how much and to what extent can the process of extrapolation be quantified when applied to an individual, and what assumptions are made when this process is undertaken?

This paper will outline how results of RCTs or meta-analyses of RCTs can be interpreted in the context of treating an individual patient. An example of a patient with acute sinusitis illustrates how this external evidence can be incorporated into an individual patient decision (Box 1).

Problems with RCTs in general practice

Asking the right question

Clearly, not all clinical problems can be solved by resorting to an RCT. By framing a clinical question in an appropriate way, the relevant methodology suitable for answering that question can be selected.[9] Much confusion concerning the limitations of RCTs, and indeed EBM, can be avoided by clarifying a clinical question before commencing a search.[3,10] Many problems in general practice (for example, attitudes and perceptions towards illness and treatment) can only be answered by qualitative means. Similarly, in some situations it may be unethical or inappropriate to perform an RCT, and quantitative observational research may provide the most suitable answer.[11]

Finding the evidence

There are many apparent obstacles to seeking out evidence in the form of RCTs and applying them to patients in general practice. A common perception is that there is a paucity of evidence on which to base clinical decisions. This view is no longer sustainable. There have been several studies that have sought to identify and categorize the published output of RCTs in general practice journals, both in the United Kingdom[12] and in the United States.[13] From these sources, a register of trials relevant to general practice has been established.[14] In addition, primary care is a recognized 'field' of interest within the Cochrane Collaboration — an acknowledgement that by its nature it is multi-disciplinary, with evidence being found from a wide range of sources. More recently, unpublished trials and RCTs published in non-primary care journals can be readily found in the Cochrane register of RCTs, part of the Cochrane Library.[15]

Setting versus patient characteristics

Another apparent barrier is that much evidence used in general practice emanates from RCTs that have been performed in different settings, making the patients unlike those seen in primary care. However, as has been observed in many conditions (e.g. hypertension, left ventricular failure, hypercholesterolaemia, non-rheumatic atrial fibrillation, and treatment of HIV), it is the patient's characteristics and not the setting that determines the risk and consequent benefit of treatment.[16-18] Discounting a treatment on the grounds that the RCT did not take place in a general practice setting is misleading. Indeed, when quantitative estimates of co-morbidity in patients usually seen in general practice are compared with the patients enrolled in RCTs (e.g. in hypertension[19]), the benefits of treatment may be even more substantial because healthier individuals tend to be over-represented in RCTs.

Representativeness of a trial population

An RCT may not be representative of a population because many eligible patients are excluded. The published reports of difficulty in recruitment tend to reinforce this view.[20-22] However, there have been many notable successes of well-conducted RCTs based in general practice.[23] The situation is likely to change: the proportion of eligible individuals eventually randomized in each RCT will be more transparent with the introduction of the CON-

T Fahey, MD, MSc, MRCGP, MFPHM, senior lecturer in primary care, Division of Primary Care, Department of Social Medicine, University of Bristol.
Submitted: 7 May 1997; accepted: 12 September 1997.

SORT guidelines in major medical journals, so individual clinicians will be able to judge for themselves the shortcomings of an RCT in terms of the representativeness of the study population.[24] When dealing with individuals, the clinical question becomes one of 'applying' rather than 'generalizing'.[25] Thus, rather than generalizing to all patients, the GP needs to consider whether the patient for whom treatment is being considered is so different from trial participants that the results of a study might somehow differ if applied to that person.

The framing of risk

Understanding and interpreting risk is crucial to rational decision making.[26] Unfortunately, it has been demonstrated that hospital doctors, GPs, or patients do not appreciate the various indices by which risk is conveyed in RCTs.[27-29] There are a variety of ways in which the risks and benefits of treatment can be presented, and these are summarized in Table 1.[30,31] Ideally, a single estimate of risk should be presented with the corresponding 95% confidence interval, thus conveying the degree of uncertainty around an estimate.[32] Whatever index of risk is chosen, pertinent to the process of rational decision making is an understanding of how risk is formulated.

Acting on the evidence

When new treatments emerge, it can take some time before clinical practice adopts new evidence.[33-35] There is some evidence to suggest that GPs are slower in adapting to change than are specialists.[36] However, all clinicians face the challenge of maintaining best practice and implementing therapies that have been shown to be effective.[37,38] Further research into the barriers to change in general practice is needed.

Quantifying treatment benefit and harm

Quantification encourages explicit decision making. All therapeutic decisions are concerned with maximizing benefit and minimizing harm.[1,26] By making the process explicit, GPs will be able to clarify several points to their patients: the outcome of interest, deferment of disability or increase in quality of life, and the likelihood of benefit and harm and their associated costs. The clinical scenario of acute sinusitis is used as an example of how this explicit process may be used (Box 1).

The advantage of this approach is that the demarcation between external evidence and clinical judgement is clarified.[10] It may also help GPs to appreciate why many 'experts' — who frequently fail to be explicit in the reasons behind their therapeutic recommendations — differ when considering the same therapeutic problem.[39]

Underlying Assumptions

Heterogeneity

Heterogeneity refers to differences between groups. It may be useful to view heterogeneity in a statistical and clinical context.[40,41] When the results of RCTs are combined by means of meta-analysis, a statistical test of heterogeneity is applied to assess whether the individual treatment estimates for each of the RCTs differs beyond what could be expected by chance.[41] Thus, the statistical test of heterogeneity should be viewed as a means of testing whether or not it is valid to combine the results of individual RCTs into a summary measure of effect. Clinical heterogeneity refers to differences in characteristics of the populations, interventions, and outcomes of RCTs that are included in a meta-analysis. Therefore, clinical heterogeneity can be usefully viewed as a difference in the design of the individual RCTs that form a meta-analysis; such differences may or may not result in statistical heterogeneity.[40]

Consequently, when critically appraising a meta-analysis, a non-significant test of heterogeneity is no guarantee that important differences in the design of individual RCTs are not present.[40] It is often far better to critically examine the possible explanations for heterogeneity — such as differences in patient characteristics (age, sex, socio-economic status, patients with established disease or without disease), interventions (timing, type and dosage of drugs, whether other therapies such as diet have also been examined), and outcomes (morbidity, mortality, quality of life) — and look upon a systematic review as a 'valuable objective descriptive technique' rather than a simple answer to a complex clinical problem.[40] Whatever simple summary statistics of effect are used, they should always be viewed and applied in the context of the complexity of the underlying design of the RCT or meta-analysis in question.

Baseline risk

For any group of individuals enrolled in an RCT, the risk of developing the outcome of interest necessarily varies among the population. Thus, the cohort of individuals in every RCT are heterogeneous in their baseline risk.[42] A frequent assumption is made that the relative benefit of treatment is proportionate across different strata of baseline risk, and indeed for most conditions this assumption holds. For example, the relative benefits of HMG CoA reductase inhibitor drugs in the treatment of hypercholesterolaemia is approximately 30% when cardiovascular and total mortality is the outcome of interest.[43,44] However, the baseline risk differs substantially both within studies (absolute risk reduction in the highest versus lowest quintiles of risk in the 4S study = 9.8% and 5.8% respectively)[44] and between studies (absolute risk reduction = 2.5% in the WOSCOP study).[43,45] The logical conclusion is that individuals at higher absolute risk yield fewer individuals who require treatment to prevent a single event.[25]

However, the assumption of constant relative benefit across different strata or subgroups of baseline risk does not necessarily hold. In some situations, the relative benefit may differ in magnitude across subgroups (so-called heterogeneity of effect); in others, treatment may be harmful or beneficial in different subgroups of patients (qualitative heterogeneity of effect).[42,46] For example, the relative benefit of treatment with aspirin compared with placebo in patients who have had a transient ischaemic attack (TIA) or minor ischaemic stroke has been shown to diminish as baseline risk increases, making the absolute benefits of treatment in patients with high baseline risk less than those individuals with low baseline risk.[47] An example of qualitative heterogeneity was demonstrated when the results of carotid surgery in patients who had already suffered a TIA was compared with placebo. Individuals with the highest baseline risk benefited from treatment, while those in the lowest baseline risk group were more likely to be harmed than helped by surgery.[47]

Finally, when estimating each individual's likely baseline risk, it is preferable to rely on independently derived prognostic models based on large cohorts of patients in which measurement of each patient's characteristics has taken place.[41,47-49] Using the observed proportion of events in the control arm as a measure of underlying risk in meta-analyses is flawed. This is because outcome in the control group is related to the treatment effect. Because the treatment effect is derived in part from the control group outcome, a relationship occurs. Such a relationship inflates estimates of treatment benefit in high-risk patients and reduces

The patient problem

Patient 1 is a 34-year-old woman who complains of malaise associated with purulent nasal discharge of one week's duration. She also complains of unilateral facial pain, toothache, and anosmia. These symptoms are exacerbated on bending forward. On examination she is pyrexial; there is obvious nasal congestion and purulent rhinorrhoea, and she is tender over her maxillary sinus on the left side. Patient 2 is a man aged 45 who complains of headache, purulent nasal discharge, and pain in his right upper teeth for two weeks, which have failed to respond to topical nasal decongestants. On examination he is apyrexial and has no clinical signs of abnormality.

The clinical problem

You are reasonably confident that both of these patients do indeed have acute sinusitis, despite the fact that it is a difficult diagnosis to make on history and examination alone.[68] The clinical question you choose to frame is as follows: 'Do patients with acute sinusitis diagnosed in a general practice setting benefit from treatment with antibiotics in terms of reducing the severity of their illness, and, if so, do the benefits of antibiotic treatment exceed their side effects?'

You perform a MEDLINE search from 1992 to 1996 combining the search terms 'clinical trial in pt (publication type)' and the exploded medical subject heading (MESH) 'sinusitis', and get 78 hits. Among these trials you find an RCT comparing penicillin V or amoxycillin to placebo, which has been performed in a Norwegian general practice involving 130 adult patients whose symptoms and signs have been validated by the use of a CT-scan to confirm acute sinusitis.[69]

Estimating benefits and side effects

On reading the study, you are happy that it is indeed a valid study and that the results are presented clearly. A summary of the benefits of antibiotic treatment is presented in Table 1, while the side effects of therapy are reported in Table 2. When explaining the results of the RCT to each of the patients, you feel you can elucidate to them three points. First, acute sinusitis is frequently a self-limiting illness; in the RCT over half of the patients (57%) got better without any further treatment. Secondly, the relative risk reduction of a 10-day course of antibiotics is substantial (RRR was 67%), and this translates into a low number of patients who need to be treated (NNT) for an acute episode of sinusitis to be prevented. Thirdly, side effects of treatment, though not usually severe (see the severity rating scale in Table 4 of the original RCT),[69] are also common: the absolute increase in side effects of treatment (ARI was 21%) was similar to the absolute benefits of treatment (ARR was 29%).

Individual baseline risk

In order to 'fine tune' the results of this RCT to each patient, you need to calculate two further pieces of information. The first is the extent to which each patient is either more or less ill than the patients enrolled in the RCT, technically estimating the likely baseline risk or patient expected event rate (PEER) for each patient.[31] This individual susceptibility has been denoted the *f* value and is the estimated amount by which individual susceptibility is greater or less than that of a typical patient enrolled in the RCT that is being applied.[25,31] For example, if the individual risk of a patient is estimated to be half that of a patient enrolled in an RCT, then *f* is equal to 0.5. Baseline risk in the trial is multiplied by the *f* factor to give an estimate of each individual's susceptibility to an adverse outcome.[25,31]

You intuitively feel that patient 1 is more unwell than patient 2. Fortunately, the RCT you have found has quantified the severity of illness according to a severity scale dependent on the presence of symptoms and signs. The scale runs from 0 to 13, with mean scores of 8.4 to 8.3 in the patients enrolled in the RCT (Table 2).[69] From this data you are able to score the severity of illness in both of your patients: patient 1 scores 13, while patient 2 scores 4. Thus, when the benefit of treatment is considered, the f value is 1.55 in patient 1, and 0.5 in patient 2. You estimate that neither patient is more or less likely to suffer side effects from antibiotic treatment than a typical patient in the RCT. Thus, the f value equals 1 for both patients when the absolute risk increase is considered.

Side effects versus benefits: patient's view

Secondly, you need to ask each patient how she or he views the relative importance of reducing the severity of their sinusitis attack (ARR) when compared with the relative increase in getting side effects from antibiotic treatment (ARI). Though side effects of therapy are relatively common, they appear to be mild (the mean severity score was between 24 and 42 out of a rating scale of 100).[69] You then ask each patient how they view prevention of their sinusitis symptoms in the context of getting side effects from antibiotics. Patient 1 views the likelihood of side effects as half as bad as continued sinusitis, while patient 2 views the risk of side effects as being the same as the risk of continued sinusitis. From these estimates it is possible to quantify the relative importance of benefit compared with the side effects of therapy for each individual patient: the *s* value. Thus, the *s* value for patient 1 is 0.5, while it is 1 for patient 2.

Likelihood of being helped or harmed (LHH)

Table 3 illustrates how benefits and side effects of treatment can be assessed by relating the likely benefit of treatment (ARR x f) to the likely harm adjusted for each patient's tolerance of side effects as a product of benefit (ARI x f x s). The resulting ratio is known as the likelihood of being helped or harmed (LHH). For the first patient, the likely benefit of antibiotic treatment outweighs side effects by a factor of 4, while for patient 2 the likely harm exceeds benefit by nearly a half. This ratio usefully conveys to patients their own preference for treatment in the context of external evidence from an RCT.

Resolution of clinical scenario

Whether overall summary measures such as NNT or LHH are used is not central to the process of applying the results of the RCT to an individual patient. As no evidence exists to demonstrate that one particular method or risk index is superior to another, clinicians should be flexible in their approach when discussing treatment risk and benefit. Indeed, the use of *s* and *f* values, particularly without use of confidence intervals, may convey inappropriate precision and certainty to individual patients. What is crucial is that a clinical problem is broken down into a meaningful question, and that the whole process of finding, appraising, and applying the external evidence is explicit and explained to the patient.[25]

Box 1. The clinical scenario.

Table 1. Summary of the results of an RCT comparing amoxycillin or penicillin V with placebo in the treatment of acute sinusitis.[69]

	Feel worse, unimproved, or somewhat better	Feel restored or much better	Total
Antibiotic	12	71	83
Placebo	19	25	44

Outcome of interest: proportion of individuals not better at day 10.

Control event rate (CER) or baseline risk = 19/44 (43%.)
Experimental event rate (EER) = 12/83 (14%).
Relative risk (RR) = ratio of antibiotic rate to placebo rate (EER/CER) = 14/43 = 0.33.
Odds ratio (OR) = (12/71)/(19/25) = 0.22.
Relative risk reduction (RRR) = (CER–EER)/CER = 67%.
Absolute risk reduction (ARR) = CER-EER = 29%.
Number of patients who need to be treated (NNT) for 10 days with an antibiotic to prevent an additional adverse event (feeling worse, unimproved, or somewhat better) = 1/ARR = 3.5 (95% CI = 2 to 8).

Table 2. Quantitative summary of the side effects of treatment in the antibiotic and placebo groups.[69]

	Feel worse, unimproved or somewhat better	Feel restored or much better	Total
Antibiotic	49	37	86[a]
Placebo	16	28	44

Outcome of interest: side effects of treatment (includes diarrhoea, nausea/vomiting, rash, vaginal discharge, and unspecified other side effects). [a]Three patients in the antibiotic group experienced side effects early on in the study and did not get any further treatment, hence they are not included in the denominator in Table 1.

Control event rate (CER) or baseline risk = 16/44 (36%).
Experimental event rate (EER) = 49/86 (57%).
Relative risk (RR) = ratio of placebo rate to antibiotic rate (CER/EER) = 36/57 = 0.63.
Absolute risk increase (ARI) = EER–CER = 21%.
Number of patients who need to be harmed (NNH) after 10 days when prescribed an antibiotic compared with placebo = 1/ARI = 5 (95% CI = 3 to 38).

Table 3. Calculation of the likelihood of being helped or harmed by treatment (LHH).

	Patient 1	Patient 2
Mean clinical severity score reported in RCT[69]	8.4	8.4
Estimated severity score in each patient calculated from RCT	13	4
f score (relative severity of illness)	1.55	0.5
Absolute risk reduction (ARR)	0.29	0.29
Absolute risk increase (ARI) of side effects of treatment	0.21	0.21
f score (relative likelihood of suffering side effects)	1	1
s score (severity of side effects in relation to benefits of treatment)	0.5	1
Likely benefit (*f* x ARR)	0.445	0.145
Likely harm (*f* x ARI x s)	0.105	0.21
Likely benefit : likely harm (LLH)	4.2 : 1	1 : 1.45

the same estimates in low-risk patients.[41,48] A shortcoming of using this approach is that relevant prognostic models are seldom available for GPs to use in the management of individual patients.[50,51]

Subgroup analyses

By identifying the subgroup of patients who, because of their comorbid clinical characteristics, are at a higher risk, treatment may be targeted at these individuals. Subgroup analyses, by their very nature, usually occur retrospectively. The danger of multiple subgroup comparisons is that a type 1 error — rejecting the null hypothesis of no difference between active treatment and placebo when in fact this is true — is more likely to occur simply by chance.[46] However, prospective trials examining potential subgroups that may benefit from treatment may not always be feasible, possible, or ethical.[52] A guide to assessing whether apparent differences in treatment benefit in subgroups are real has been proposed.[46] Before applying results of an RCT or meta-analysis to an individual patient on the strength of a sub-group analysis, it is advisable to apply such criteria.[46]

Tolerability

Rates of discontinuation when taking drugs for chronic disease (e.g. hypertension and hypercholesterolaemia) have been examined in cohorts of patients based in the community. These studies have demonstrated that rates of discontinuation (used as a proxy measure for tolerability) have been higher than would be expected from the reported side effects of treatment in patients enrolled in RCTs.[53,54] Many RCTs select patients who are likely to be compliant so that the efficacy of treatment can be examined in ideal circumstances.[55,56] Caution is required in assuming that individuals in the community necessarily have the same tolerance to side effects as individuals selected in trials. This relates to whether the RCT on which a therapeutic decision is to be made is pragmatic or explanatory in nature.[56] Little information is known about tolerance to side effects in community settings outside large RCTs. Large community-based prescribing databases may provide more information in the future. Until more is known it should not be assumed that side effects reported in RCTs reflect, with certainty, actual rates in the community.

Future approaches

Explicit quantification of risk and benefit, as laid out in the clinical scenario, is in its infancy, and the factors of *f* and *s* used in the accompanying clinical scenario (Box 1) are attempts to measure what is viewed as clinical judgement. Clinical judgement and knowledge of people play a crucial, additional role in the application of external evidence from RCTs.[57] Many barriers against adopting such an approach currently exist in general practice, for example paucity of information, diagnostic uncertainty, and time constraints during the consultation. Indeed, effective methods of implementation rely on a multi-faceted approach using diverse interventions such as reminders, patient-mediated interventions, outreach visits, and opinion leaders.[58] In particular, the use of computers in the form of clinical-based decision support systems (CDSSs) has not been fully exploited as a means of quantifying risk and benefit, and could play an important role in the implementation of evidence during consultations in the future.[59] Furthermore, an understanding and appreciation of patient preferences is likely to play a greater role in clinical decision making.[60] For example, patients with a similar severity of angina on a clinical rating scale varied considerably in the tolerance of their symptoms when measured according to a utility score.[61] Decision analysis, which combines probability of illness with individual utility, is likely to provide a more robust approach to combined decision making.[26]

Not all treatments will be amenable to measurement by means of an RCT. There will be areas where evidence is either incomplete or conflicting. In such situations it becomes even more

important to acknowledge doubt and to respect patient preferences.[62] For patients in general practice, problems are seldom one-dimensional, and implementation of research may require a search for descriptive and qualitative studies.[63] Acknowledgement of uncertainty, respect for patient preferences, and an ability to clarify clinical problems will require understanding and flexibility in all GPs who wish to do the best for their patients.

Conclusion

Efficient searching for relevant information and keeping up to date with new research findings remains a constant challenge for all doctors.[25] While this article was being written, another RCT suggesting that antibiotics have little effect on the course of acute sinusitis was published.[64] This highlights the need for caution when applying the results of individual RCTs, and emphasizes the strength of systematic reviews that aim to identify and update all new evidence.[65] The totality of evidence changes all the time and requires constant re-evaluation from individual GPs.[65] The Cochrane Collaboration (through its register of RCTs and database of systematic reviews) and other similar initiatives should help individual clinicians in finding the best evidence.

When applying the results of RCTs or meta-analyses to individual patients, critical appraisal of the study is right and necessary. However, it is always worth considering at the start what is known and not known about a particular treatment. Using external evidence in the form of an RCT enables an informed decision to take place with each patient in the context of the severity of their illness. When using indices of risk, benefit, and harm, it is best that absolute and relative measures should be used.[30] Finally, patients should be encouraged to make decisions about their treatment on the basis of the most complete evidence.[61,66] As Holmberg and Baum state, 'a critical discussion will be most fruitful if we openly accept that RCTs are powerful instruments in setting limitations on our ignorance.'[67]

References

1. Sackett DL, Haynes RB, Guyatt GH, Tugwell P (eds). *Clinical epidemiology: A basic science for clinical medicine.* 2nd edn. Toronto: Little, Brown and Company, 1991.
2. The evidence based medicine working group. Evidence based medicine: a new approach to teaching the practice of medicine. *JAMA* 1992; **268:** 2420-2425.
3. Sackett DS, Rosenberg W, Gray JAM, *et al.* Evidence based medicine: what it is and what it isn't. *BMJ* 1996; **312:** 71-72.
4. Dawes MD. On the need for evidence-based general and family practice. *Evidence-Based Medicine* 1996; **1:** 68-69.
5. Hennekens C, Buring J. *Epidemiology in medicine.* Boston/Toronto: Little, Brown and Company, 1987.
6. Guyatt GH, Sackett DL, Sinclair JC, *et al.* Users' guides to the medical literature IX. A method for grading health care recommendations. *JAMA* 1995; **274:** 1800-1804.
7. Chalmers I, Enkin M, Keirse MJNC. Preparing and updating systematic reviews of randomised controlled trials of health care. *The Milbank Quarterly* 1993; **71:** 411-437.
8. Channer K. Translating clinical trials into practice. *Lancet* 1997; **349:** 654.
9. Oxman A, Sackett D, Guyatt G. Users' guides to the medical literature. I. How to get started. *JAMA* 1993; **270:** 2093-2095.
10. Sullivan F, MacNaughton R. Evidence in consultations: interpreted and individualised. *Lancet* 1996; **348:** 941-943.
11. Pringle M, Churchill R. Randomised controlled trials in general practice. *BMJ* 1995; **311:** 1382-1383.
12. Silagy C, Jewell D. Review of 39 years of randomized controlled trials in the British Journal of General Practice. *Br J Gen Pract* 1994; **44:** 359-363.
13. Silagy CA, Jewell D, Mant D. An analysis of randomized controlled trials published in the US family medicine literature, 1987-1991. *J Fam Pract* 1994; **39:** 236-242.
14. Silagy CA. Developing a register of randomised controlled trials in primary care. *BMJ* 1993; **306:** 897-900.
15. Cochrane Collaboration. *The Cochrane Library. V (volume 5?).* Issue 2. Oxford: Update Software, 1997.
16. Davey Smith G, Egger M. Who benefits from medical interventions? *BMJ* 1994; **308:** 72-74.
17. Davey Smith G, Song J, Sheldon T. Cholesterol lowering and mortality: the importance of considering initial level of risk. *BMJ* 1993; **306:** 1367-1373.
18. White HD. Aspirin or warfarin for non-rheumatic atrial fibrillation? *Lancet* 1994; **343:** 683-684.
19. Mulrow CD, Cornell JA, Herrera CR, *et al.* Hypertension in the elderly: implications and generalizability of randomised trials. *JAMA* 1994; **272:** 1932-1938.
20. Tognoni G, Avanzini F, Bettelli G, *et al.* Randomised clinical trials in general practice: lessons from a failure. *BMJ* 1991; **303:** 969-971.
21. Peto R, Coulter A, Bond A. Factors affecting general practitioners' recruitment of patients into a prospective study. *Fam Pract* 1993; **10:** 207-211.
22. Thomas S. Antibiotics for cough and purulent sputum. *BMJ* 1978; **i:** 1374.
23. Hart JT. Randomised controlled trials in general practice: General practice framework carries out such trials. *BMJ* 1996; **312:** 779.
24. Altman D. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996; **313:** 570-571.
25. Sackett D, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based medicine: How to practice and teach EBM.* London: Churchill Livingstone, 1996.
26. Sox H, Blatt M, Higgins M, Marton K. *Medical decision making.* Boston: Butterworths, 1988.
27. Laupacis A, Sackett DL, Roberts RS. Therapeutic priorities of Canadian internists. *Can Med Assoc J* 1990; **142:** 329-333.
28. Cranney M, Walley T. Same information, different decisions: the influence of evidence on the management of hypertension in the elderly. *Br J Gen Pract* 1997; **46:** 661-663.
29. Hux J, Levinton C, Naylor C. Prescribing propensity: influence on life-expectancy gains and drug costs. *J Gen Intern Med* 1994; **9:** 195-201.
30. Sackett D, Cook R. Understanding clinical trials. *BMJ* 1994; **309:** 755-756.
31. Cook R, Sackett D. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995; **310:** 452-454.
32. Altman DG. Use of confidence intervals to indicate uncertainty in research findings. *Evidence-Based Medicine* 1996; **1:** 102-104.
33. Ketley D, Woods KL. Impact of clinical trials on clinical practice: example of thrombolysis for acute myocardial infarction. *Lancet* 1993; **342:** 891-894.
34. The European Secondary Prevention Study Group. Translation of the clinical trials into practice: a European population based study of the use of thrombolysis for acute myocardial infarction. *Lancet* 1996; **347:** 1203-1207.
35. Lamas GA, Pfeffer MA, Hamm P, *et al.* Do the results of randomised clinical trials of cardiovascular drugs influence clinical practice. *N Eng J Med* 1992; **327:** 241-247.
36. Redelmeier DA, Tversky A. Discrepancy between medical decisions for individual patients and for groups. *N Eng J Med* 1990; **322:** 1162-1164.
37. Haines R, Jones R. Implementing findings of research. *BMJ* 1994; **308:** 1488-1492.
38. Haines A. The science of perpetual change. *Br J Gen Pract* 1996; **46:** 115-119.
39. Richards JP. Evidence based general practice. *BMJ* 1997; **314:** 525.
40. Thompson SG, Pocock SJ. Can meta-analyses be trusted? *Lancet* 1991; **338:** 1127-1130.
41. Sharp SJ, Thompson SG, Altman DG. The relation between treatment benefit and underlying risk in meta-analysis. *BMJ* 1996; **313:** 735-738.
42. Bailey K. Generalizing the results of randomised clinical trials. *Controlled Clinical Trials* 1994; **15:** 15-23.
43. Shepherd J, Cobbe SM, Ford I, *et al.* Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Eng J Med* 1995; **333:** 1301-1307.
44. Scandinavian Simvastatin Survival Study Group. Baseline serum cholesterol and treatment effect in the Scandinavian simvastatin survival study. *Lancet* 1995; **345:** 1274-1275.
45. Vandenbroucke JP, Westendorp RGJ. Are the effects of cholesterol lowering drugs always equal? *Lancet* 1996; **347:** 1267-1268.
46. Oxman A, Guyatt G. A consumer's guide to subgroup analysis. *Ann Intern Med* 1992; **116:** 78-84.
47. Rothwell PM. Can overall results of clinical trials be applied to all patients? *Lancet* 1995; **345:** 1616-1619.
48. Egger M, Davey Smith G. Risks and benefits of treating mild hypertension: a misleading meta-analysis? *J Hypertens* 1995; **13:** 813-815.

49. Glasziou P, Irwig L. An evidence based approach to individualising treatment. *BMJ* 1995; **311:** 1356-1359.
50. Shaper A, Pocock S, Phillips A, Walker M. Identifying men at high risk of heart attacks: strategy for use in general practice. *BMJ* 1986; **293:** 474-479.
51. Coppola W, Whincup P, Papacosta O, *et al.* Scoring system to identify men at high risk of stroke: a strategy for general practice. *Br J Gen Pract* 1995; **45:** 185-189.
52. Sørensen TIA. Which patients may be harmed by good treatment? *Lancet* 1996; **348:** 351-352.
53. Jones J, Gorkin L, Lian J, *et al.* Discontinuation of and changes in treatment after start of new courses of antihypertensive drugs: a study of a United Kingdom population. *BMJ* 1995; **311:** 293-295.
54. Andrade SE, Walker AM, Gottlieb LK, *et al.* Discontinuation of antihyperlipidaemic drugs - do rates reported in clinical trials reflect rates in primary care settings? *N Eng J Med* 1995; **332:** 1125-1131.
55. Bulpitt CJ, Fletcher AE. Compliance with antihypertensive treatment in the elderly. *Handbook of hypertension* 1989; **12:** 387-397.
56. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in the therapeutical trials. *J Chron Dis* 1967; **20:** 637-648.
57. McCormick JS. The place of judgement in medicine. *Br J Gen Pract* 1994; **44:** 50-51.
58. Davis D, Thomson M, Oxman A, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. *JAMA* 1995; **274:** 700-705.
59. Johnson M, Langton K, Haynes R, Mathieu A. Effects of computer-based clinical decision support systems on clinician performance and patient outcome. *Ann Intern Med* 1994; **120:** 135-142.
60. Redelmeier DA, Rozin P, Kahneman D. Understanding patients' decisions. *JAMA* 1993; **270:** 72-76.
61. Nease R, Kneeland T, O'Connor G, *et al.* Variations in patient utilities for outcomes of the management of chronic stable angina. *JAMA* 1995; **273:** 1185-1190.
62. Naylor CD. Grey zones of clinical practice: some limits of evidence-bases medicine. *Lancet* 1995; **345:** 840-842.
63. Greenhalgh T. Is my practice evidence-based? *BMJ* 1996; **313:** 957-958.
64. Van Buchem FL, Knottnerus JA, Schrijnemaekers V, Peeters MF. Primary-care based randomised placebo-controlled trial of antibiotic treatment in acute maxillary sinusitis. *Lancet* 1997; **349:** 683-687.
65. Mulrow CD. Rationale for systematic reviews. In: Chalmers I, Altman D (eds). *Systematic reviews*. London: BMJ Publishing Group, 1995.
66. Kassirer JP. Incorporating patients' preferences into medical decisions. *N Eng J Med* 1994; **330:** 1895-1896.
67. Holmberg L, Baum M. Can results from clinical trials be generalised? *Nature Medicine* 1995; **1:** 734-736.
68. Williams JW, Simel DL. Does this patient have sinusitis? *JAMA* 1993; **270:** 1242-1246.
69. Lindbaek M, Hjortdahl P, Johnsen UL-H. Randomised, double blind, placebo controlled trial of penicillin V and amoxycillin in treatment of acute sinus infections in adults. *BMJ* 1996; **313:** 325-329.

Acknowledgements

I would like to thank Deirdre Murphy, Yoav Ben Schlomo, George Davey Smith, and Debbie Sharp for constructive criticisms on early drafts of this paper, and the Evidence-Based Medicine discussion group, the original sources concerning the *s* and *f* values when applied to individual patients.

Address for correspondence

Dr T Fahey, Division of Primary Care, Department of Social Medicine, University of Bristol, Canynge Hall, Whiteladies Road, Bristol BS8 2PR. Email: tom.fahey@bris.ac.uk