An evidence based approach to individualising treatment

Paul P Glasziou, Les M Irwig

To which groups of patients can the results of clinical trials be applied? This question is often inappropriately answered by reference to the trial entry criteria. Instead, the benefit and harm (adverse events, discomfort of treatment, etc) of treatment could be assessed separately for individual patients. Patients at greatest risk of a disease will have the greatest net benefit as benefit to patients usually increases with risk while harm remains comparatively fixed. To assess net benefit, the relative risks should come from (a meta-analysis of) randomised trials; the risk in individual patients should come from multivariate risk equations derived from cohort studies. However, before making firm conclusions, the assumptions of fixed adverse effects and constant reduction in relative risk need to be checked.

Should all patients with acute myocardial infarct receive streptokinase? Should all patients with non-valvar atrial fibrillation receive warfarin? Such questions are best answered by assessing benefits and risks in each patient rather than focusing on the inclusion and exclusion criteria of the trial. For example, patients with a history of peptic ulcer were often excluded from thrombolytic trials because it would be unethical to put such patients at risk of major bleeding episodes, given the lack of proved benefit. However, now that we know the size of the benefit—about three deaths saved for every 100 patients treated[1]—clinicians must make an informed decision that weighs this benefit against the potential for harm in patients with ulcers.

Relying on the eligibility criteria for clinical trials is both erroneous and limiting. A too restrictive generalisation needs to be guarded against, and we are advised to ask, "Are the patients in this study so different from my patients that I could not apply the study results?"[2] This is good advice, but how then do we decide when a patient is too different to benefit from treatment? The search for differences should be based on features of the disease process or risk rather than differences in sociodemographic characteristics.[3]

The converse of this question is, "Can the study results be generalised to all patients who would be eligible for the trial?" The answer might seem to be obviously "yes." However, we will show that this is also incorrect after we have developed a general approach.

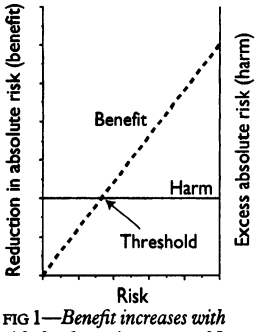

FIG 1—*Benefit increases with risk, but harm is constant. Net benefit occurs only when risk is above threshold*

The basic model: separating benefit and harm

Lubsen and Tijssen proposed a separate assessment of the benefit and harm of treatment.[4] As shown in figure 1, their model suggests patient benefit increases with risk from the disease—those most at risk have most to gain—but that harm or rates of adverse event will remain comparatively fixed. Thus at some low level of risk the benefits will only just balance the harm and we should refrain from treatment.[5,6]

This model works by converting the reduction in relative risk, which is useful for assessing the strength of the intervention, to a reduction in absolute risk, which is useful for assessing the clinical worth of the intervention. The relative risk is the ratio of clinical events in the treated group relative to the control group. For example, if the control group had a death rate of 12% and the treated group a death rate of 9% the relative risk is 9/12, or 75%, which implies a 25% reduction in relative risk with treatment. However, the reduction in absolute risk is 12%−9%, or 3%.

How do we extrapolate for patients at different risk? The reduction in absolute risk of 3% is unlikely to apply universally. For example, if the control group had a death rate of only 2% instead of 12% a reduction in absolute risk of 3% would give the impossible death rate of 2%−3%, or −1%. It is usually more reasonable to assume that the reduction in relative risk stays constant, which would suggest that the treated group would have a death rate of 1·5% (still a 25% reduction in relative risk but only a 0·5% reduction in absolute risk).

The model of Lubsen and Tijssen extends the above calculations to incorporate a fixed harm. This is illustrated in figure 2 and may be expressed in the following equation:

Net benefit=risk×reduction in relative risk−harm.

To apply this model requires four steps.

(1) Estimate benefit and harm

A randomised trial, or a meta-analysis of randomised trials, is the most appropriate method to estimate the reduction in relative risk with the intervention for various outcomes relevant to patients. Methods to identify, appraise, and combine trials are described elsewhere.[7] Such an analysis is vital for proving an effect and estimating its comparative size. This is, however, only a first step in deciding which patients would expect to benefit from a treatment.

EXAMPLE: ANTICOAGULANTS AND NON-RHEUMATIC ATRIAL FIBRILLATION

There have been six randomised controlled trials of low dose warfarin in a total of 4269 patients with non-rheumatic atrial fibrillation. Five of these trials have

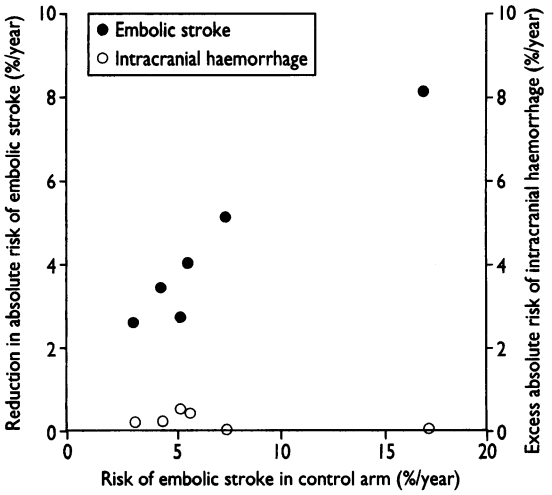

FIG 2—*Trials of warfarin in non-valvar atrial fibrillation show that benefit (reduction in absolute risk of stroke) increases with increasing risk of stroke but that harm (intracranial haemorrhage) seems to be constant*

Department of Social and Preventive Medicine, Medical School, Herston, Queensland, Australia 4006
Paul P Glasziou, *senior lecturer in clinical epidemiology*

Department of Public Health and Community Medicine, Building A27, University of Sydney, Sydney, New South Wales, Australia 2006
Les M Irwig, *professor in epidemiology*

Correspondence to:
Dr Glasziou.

BMJ 1995;311:1356-9

recently been reviewed by Singer, who found an overall reduction in relative risk of about 73%.[8] The table describes these five trials, as summarised by Singer, plus the more recent European atrial fibrillation trial.[9]

The reduction in the risk of stroke is dramatic. However, clinicians have been cautious because of the incidence of bleeding induced by warfarin, particularly the risk of intracranial haemorrhage.

(2) Check assumptions of relative benefit and absolute harm

The model assumes that the reduction in relative risk for benefit stays the same for all risks. This will not always be true. Hence a necessary check is a meta-analytic review of all trial data, with an examination of whether the reduction in relative risk varies with risk. Only if this seems to be constant should we calculate a combined estimate. Similarly, we need to check that the absolute harm is independent of risk.

Several factors may cause the reduction in relative risk to vary. Some diseases are more than one disease process, each of which may respond differently to the treatment. For example, a stroke can be embolic, thrombotic, or haemorrhagic. If stroke is considered to be a single entity, preventive treatment such as anticoagulant or antiplatelet drugs may benefit one group (with few haemorrhagic strokes) but cause net harm to another group (with mostly haemorrhagic strokes), although the risk is similar.

EXAMPLE

The reductions in relative risk in the table are reasonably constant for stroke rates in the control groups. Figure 2 plots the reduction in absolute risk for both the benefit—that is, the reduction in embolic strokes—and the harm—that is, the rates of intracranial haemorrhage. This shows that the reduction in absolute risk of thromboembolic strokes rises linearly with the risk of stroke, suggesting that the model of reduction in relative risk is appropriate. In addition, the rate of intracranial haemorrhage seems to be stable across varying risks of stroke. Thus both of the necessary assumptions seem to be fulfilled.

The assumption of a constant reduction in relative risk may sometimes be violated when the intervention has both positive and negative effects on one outcome. For example, Boissel *et al* have performed a meta-analysis on data from 13 published trials of class I antiarrhythmic agents after myocardial infarction.[10] The reduction in relative risk clearly varied with the mortality in the control group. Patients at low risk had a relative increase in mortality, presumably because of the proarrhythmic effects of these agents. However, patients at high risk (mortality greater than 15% per year) showed a net beneficial effect. Thus the assumption that the reduction in relative risk is constant is not fulfilled. Boissel *et al* then modelled this as a constant proarrhythmic harm and a relative antiarrhythmic benefit. Here the harm and benefit cannot easily be separated because clinical outcome is the same for both, and the only way to disentangle the benefit and harm is to do a very large trial(s) with a wide range of risk groups.[10]

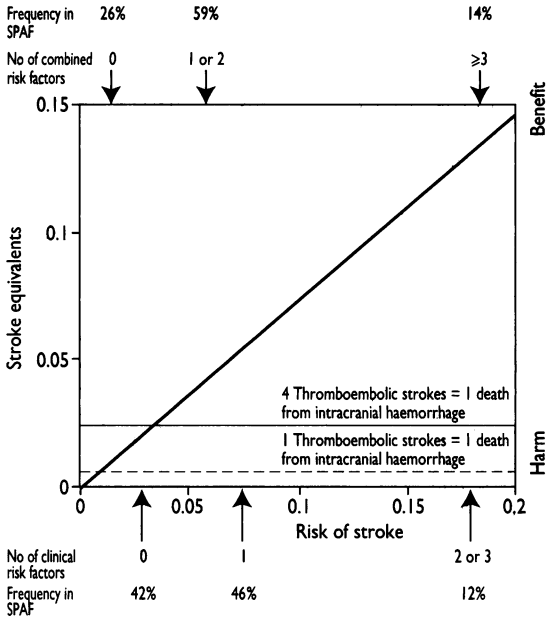

FIG 3—*Benefit compared with harm for warfarin on basis of six trials. If one death from an intracranial haemorrhage is equivalent to four thromboembolic strokes, threshold is about 2% per year. This may be predicted from three clinical risk factors (previous embolism, recent cardiac failure, and hypertension; bottom axis) or three clinical plus two echocardiographic risk factors (clinical risk factors plus atrial size and left ventricular dysfunction; bottom axis). SPAF=stroke prevention in atrial fibrillation trial*

Factors that may modify the relative reduction in risk achieved by an intervention, such as the intensity of the intervention, the timing and mode of administration, and the risk, should be checked before moving on to the next step. This can be done graphically, as we have done for risk in figure 2, and statistically.[11-13]

(3) Weigh up benefit and harm

If the assumptions of relative risk reduction and constant harm are fulfilled the predicted benefit then needs to be weighed up against the potential harm.

EXAMPLE

If we now use the estimates of the reduction in relative risk for thromboembolic stroke[8] and the risk for fatal haemorrhage[14] from meta-analyses in the equation for net benefit we obtain figure 3. To choose the point where benefit outweighs harm we need an equivalence between the thromboembolic strokes and fatal intracranial haemorrhages (within the six randomised trials all patients with an intracranial

Results from six randomised controlled trials of warfarin for non-valvar atrial fibrillation[8,9]

	Boston area anticoagulation trial for atrial fibrillation	Veterans Affairs stroke prevention in atrial fibrillation study	Canadian atrial fibrillation anticoagulation study	Atrial fibrillation, aspirin, anticoagulation study	Stroke prevention in atrial fibrillation	European atrial fibrillation trial
Warfarin group:						
No of emboli/No of subjects	2/212	4/260	5/187	4/335	6/210	20/225
No of person years	487	456	200	250	260	507
Annual rate (%)	0·41	0·88	2·5	1·6	2·3	4
Control group:						
No of emboli/No of subjects	13/208	19/265	11/191	21/336	18/211	50/214
No of person years	435	440	212	373	244	405
Annual rate (%)	3·0	4·3	5·2	5·6	7·4	12
% Reduction in relative risk (95% confidence interval)	86 (51 to 96)	79 (52 to 90)	52 (−36 to 87)	71 (23 to 90)	69 (27 to 85)	66 (43 to 80)
Intracranial haemorrhage due to warfarin (%)	0·21	0·22	0·50	0·40	0	0

haemorrhage who took warfarin died). Thus to balance the benefit and harm, we can ask how many thromboembolic strokes are equivalent to one death. Recent measurements suggest that the average quality of life after thromboembolic stroke is between 0·7 and 0·8 on a scale of 0 (death) to 1 (normal good health),[15] and thus the ratio is about 4 to 1. This is the higher of the two lines for harm in figure 3. Thus when the risk of stroke is somewhere between 2% and 3% per year, the harm and benefit are about equal. For higher rates than this the benefits would start to outweigh the potential harm induced by anticoagulant treatment.

The precise threshold clearly depends on the relative value patients place on thromboembolic stroke compared with death. We have used an average value, but individual preferences may need to be considered.

(4) Predict patient's risk

To identify patients who should expect benefit to be greater than harm, we need to predict each patient's risk. This requires identification of the major risk factors, and ideally their joint estimation to establish risk—that is, multivariate risk prediction. The major requirement for such prognostic studies is that a large inception cohort has been followed up for a sufficient time to predict accurately the joint effects of the risk factors.[16] This information may come from population based cohort studies or from the controlled trials themselves. However, because the eligibility criteria for trials often narrow the range of risk and the consent process may result in a somewhat different average risk, population based cohort studies are preferable.

EXAMPLE

Factors that influence the risk of stroke in atrial fibrillation include age and clinical and echocardiographic evidence of coexisting cardiovascular disease. The combined effect of these factors was examined by the Stroke Prevention in Atrial Fibrillation Investigators in a multivariate risk model, which suggested that three clinical features—hypertension, recent congestive cardiac failure, and previous thromboembolism—and two echocardiographic features—left ventricular dysfunction and atrial size—are important.[17-18] The arrows in figure 3 show the resulting rates of thromboembolic stroke for the clinical features (bottom axis) and the combined clinical and echocardiographic features (top axis). For patients with none of the three clinical risk factors (no risk factors on bottom axis) benefit and harm are about equal and hence we should decide not to treat. This would include about 42% of the patients in the stroke prevention in atrial fibrillation trial. For patients with none of either of the clinical or echocardiographic risk factors (no risk factors on top axis) the benefit clearly does not outweigh the potential harm. This would still include about 26% of patients in the stroke prevention in atrial fibrillation trial, who, in retrospect, we would suggest were not gaining a net benefit, despite being eligible for a trial with a strongly positive outcome.

In a similar vein, we might look at the predictors of harm—that is, the risk of intracranial haemorrhage with warfarin. However, the strongest predictor is the degree of stability of the prothrombin time.[19] Since this can be known only after warfarin treatment has been started we have not included it in the model, but in principle, predictors of harm can and should also be routinely incorporated.

Discussion

With current interest in evidence based medicine, more clinical decisions will be based on the results of clinical trials. A clinician deciding whether a trial's results apply to a particular patient should not focus on the inclusion and exclusion criteria of the trial—which are usually designed for improving the power of the study or maximising safety—but should try to predict whether each patient would benefit.[3] As shown in figure 4, the decision entails piecing together three types of information about benefit and harm: the reduction in relative risk, the risk, and the relative valuation of the outcomes. Though our example has used only two outcomes—stroke and death—the equation of figure 4 generalises this to multiple outcomes.

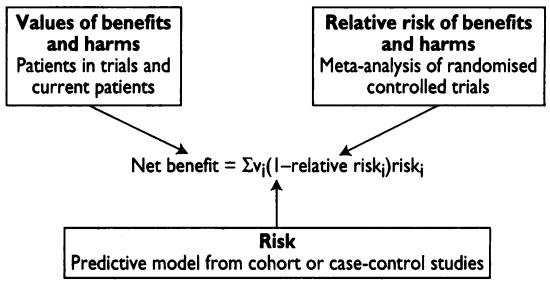

FIG 4—*Generalising clinical trials requires combining data on intervention's effectiveness with prediction of risk and patients' values. Formula is generalised version of equation in text, where v is the patient's value for outcome i relative to other outcomes, 1−relative risk$_i$ is reduction in relative risk for outcome$_i$, and risk$_i$ is predicted risk*

This method suggests that the application of trial results need not be confined by eligibility criteria or the trial's setting. Primary care settings, when compared with secondary or tertiary referral centres, are likely to have a larger proportion of patients with few or no risk factors, who would therefore enjoy little or no net benefit. In making this decision, however, the patient's specific characteristics, rather than the setting, are important.

Some readers may prefer to recast this process using the number needed to treat,[20] which is the inverse of the reduction in absolute risk, or 1/reduction in absolute risk. This would transform the vertical scale in figures 1 to 3, but the concepts and steps are unchanged.

We have examined only models in which outcome is the risk of an event. Clearly similar models could be developed for conditions in which severity is the major concern. In all cases benefit should be weighed against harm rather than relying on the eligibility criteria or setting of trials.

We thank Professor David Sackett for helpful comments.

1 ISIS-2 Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected myocardial infarction: ISIS-2. *Lancet* 1988;ii:349-60.
2 Department of Clinical Epidemiology and Biostatistics McMaster University. How to read clinical journals. V. To distinguish useful from useless or even harmful therapy. *Can Med Assoc J* 1981;**124**:1156-62.
3 Davey-Smith G, Egger M. Who benefits from medical interventions? Treating low risk patients can be a high risk strategy. *BMJ* 1994;**308**:72-4.
4 Lubsen J, Tijssen JGP. Large trials with simple protocols: indications and contraindications. *Controlled Clinical Trials* 1989;**10**:151-60S.
5 Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* 1980;**302**:1109-17.
6 Kassirer JP, Pauker SG. The toss up. *N Engl J Med* 1981;**305**:1467-9.
7 Dickersin K, Berlin JA. Meta-analysis: state of the science. *Epidemiol Rev* 1992;**14**:154-76.
8 Singer DE. Overview of the randomized trials to prevent stroke in atrial fibrillation. *Ann Epidemiol* 1993;**3**:563-7.
9 EAFT (European Atrial Fibrillation Trial) Study Group. Secondary prevention in non-rheumatic atrial fibrillation after transient ischaemic attack or minor stroke. *Lancet* 1993;**342**:1255-62.
10 Boissel J, Collet J, Lievre M, Girard P. An effect model for the assessment of drug benefit: example of antiarrhythmic drugs in postmyocardial infarction patients. *J Cardiovasc Pharmacol* 1993;**22**:356-63.
11 Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiologic Rev* 1987;**9**:1-30.
12 Maclure M. Demonstration of deductive meta-analysis: ethanol intake and risk of myocardial infarction. *Epidemiol Rev* 1993;**15**:328-51.
13 McIntosh M. The population risk as an explanatory variable in research synthesis of clinical trials. *Harvard Medical School Technical Report* 1995; No 4:1-35.

14 Landefeld CS, Beyth RJ. Risk for anticoagulant-related bleeding: a meta-analysis. *Am J Med* 1993;95:315-28.
15 Glasziou PP, Bromwich S, Simes RJ. Quality of life six months after myocardial infarction treated with thrombolytic therapy. *Med J Aust* 1994;161:532-6.
16 Laupacis A, Wells G, Richardson WS, Tugwell P, for the Evidence-Based Medicine Working Group. User's guides to the medical literature. V. How to use an article about prognosis. *JAMA* 1994;272:234-7.
17 Stroke Prevention in Atrial Fibrillation Investigators. Predictors of thrombo-embolism in atrial fibrillation. I. Clinical features of patients at risk. *Ann Intern Med* 1992;116:1-5.
18 Stroke Prevention in Atrial Fibrillation Investigators. Predictors of thrombo-embolism in atrial fibrillation. II. Echocardiographic features of patients at risk. *Ann Intern Med* 1992;116:6-12.
19 Fihn SD, McDonell M, Martin D, Henikoff J, Vermes D, Kent D, *et al.* Risk factors for complications of chronic anticoagulation: a multicenter study. *Ann Intern Med* 1993;118:511-20.
20 Laupacis A, Sackett D, Roberts R. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;318:1728-33.

(Accepted 20 August 1995)

Lesson of the Week

Occult intracranial tumours masquerading as early onset anorexia nervosa

C J DeVile, R Sufraz, B D Lask, R Stanhope

Any boy with apparent anorexia should have careful anthropometric and endocrine assessment as well as cranial imaging, which may need to be repeated

Childhood onset anorexia nervosa may be difficult to diagnose because of the lack of clear diagnostic criteria for prepubertal children and because of the difficulty of differentiating psychogenic from organic disease.[1] Psychological disturbance and symptoms of anorexia without neurological manifestation may also be presenting features of an early intracranial lesion, usually affecting the diencephalon—as shown by the following three cases.

Case reports

CASE 1

An 8½ year old boy presented with a six month history of intermittent headaches and vomiting. Clinical examination was unremarkable. A provisional diagnosis of migraine was made after unenhanced computed tomography of the brain showed nothing abnormal. Seven months after presentation the initial symptoms had settled but the patient had developed anorexia. Onset of physical symptoms coincided with the patient becoming introverted, emotionally labile, and periodically morose. His mother had died from a hepatocellular carcinoma when he was 3 years. The combination of unresolved grief, his father being admitted to hospital for a routine operation, and his eldest sister leaving home were thought to have precipitated this change in personality. He could not be coaxed into eating and was referred to a child psychiatrist, who diagnosed childhood onset anorexia nervosa with depression. Over the next nine months the patient continued solely under psychiatric follow up but with negligible improvement in weight gain or stature. His home environment was thought to be contributory, and, in view of these concerns, he was placed on the child protection register.

The patient was subsequently admitted to a second psychiatric unit for re-evaluation. Although clinically depressed, he was unresponsive to three months of psychotherapy. An endocrine opinion confirmed growth arrest both in terms of stature and weight. Despite having prolonged "anorexia," he had normal skinfold thicknesses. Formal pituitary function tests showed isolated growth hormone deficiency and a high serum prolactin concentration of 841 mU/l (normal $\leq$ 350 mU/l). Magnetic resonance imaging of the brain showed a craniopharyngioma. This was successfully resected, and the patient made an excellent recovery with immediate improvement in his mental state and resolution of his anorexia.

CASE 2

A 13 year old boy presented with a six month history of anorexia, nausea, difficulty in swallowing, and associated weight loss. Clinical examination and investigations, including barium studies, were normal. Both the patient's mother and maternal aunt had a history of anorexia nervosa in their adolescent and early adult years. The patient's mother had had symptoms for four years but eventually recovered when she met her future spouse. Although the boy had no apparent fear of fatness or distortion of body image, childhood onset anorexia nervosa was diagnosed after he had been referred to a child psychiatrist. He then underwent a stringent behaviour modification programme in hospital. During this period he was noted to have persistently dilated pupils, but visual acuity and peripheral fields were normal. There was, however, transient swelling of the optic discs. Unenhanced computed tomography of the brain showed only calcification of the pineal gland.

Nine months later the patient's condition deteriorated: he was extremely lethargic; he refused to eat, resulting in further weight loss; and he would vomit effortlessly, usually around meal times. He had also developed a right divergent squint but was not complaining of any other visual symptoms.

Psychiatric re-evaluation 10 months after presentation showed no evidence of a primary psychiatric disorder. Visual assessment confirmed a variable right divergent squint due to blindness in that eye and associated optic atrophy. Endocrine evaluation showed statural growth arrest, cortisol insufficiency, biochemical diabetes insipidus, and a high serum prolactin concentration of 7800 mU/l. Magnetic resonance imaging showed extensive enhancement of the optic nerves and chiasm, the hypothalamic area, the lower medulla, and upper spinal cord, as well as periventricularly. There was also considerable enhancement of the pineal gland. Diagnosis of a disseminating pineal germinoma was confirmed by stereotactic brain biopsy. The patient had an excellent clinical and radiological response to craniospinal irradiation and gained 12 kg in the six months after treatment.

CASE 3

A 7 year old boy was referred to our hospital with a three year history of progressive anorexia, recurrent episodes of vomiting, coughing, and discomfort on swallowing, and subsequent failure to thrive. Onset of symptoms had coincided with his father leaving home and his stepfather moving in with the family. Extensive investigation showed no organic cause for his symptoms, and computed tomography showed no abnormality, although there was no imaging below the level of the foramen magnum. Although his symptoms were atypical for anorexia nervosa, the patient was thought to have a psychological eating disorder and received

Medical Unit and Department of Psychological Medicine, Institute of Child Health and Great Ormond Street Hospital for Children NHS Trust, London WC1N 1EH
C J DeVile, *clinical research fellow*
R Sufraz, *registrar in psychiatry*
B D Lask, *consultant psychiatrist*
R Stanhope, *consultant paediatric endocrinologist*

Correspondence to:
Dr Stanhope.

BMJ 1995;311:1359-60