Personal paper

Gene therapy for cancer—managing expectations

Alan A Melcher, Ignacio Garcia-Ribas, Richard G Vile

Imperial Cancer Research Fund Molecular Oncology Unit, Royal Postgraduate Medical School, Hammersmith Hospital, London W12 0NN

Alan A Melcher, *MRC clinical training fellow*

Richard G Vile, *laboratory head*

Imperial Cancer Research Fund/ Richard Dimbleby Department of Cancer Research, St Thomas's Hospital, London SE1 7EH

Ignacio Garcia-Ribas *clinical research fellow*

Correspondence to: Dr Vile

BMJ 1997;315:1604–7

Gene therapy seems to offer new hope in cancer treatment. The new molecular technology can be used to target tumour cells in many ways. These include techniques that correct genes directly—for example, by delivering a nucleic acid sequence that complements and therefore inactivates an oncogene (antisense technology) or by replacing copies of tumour suppressor genes that are often lost in malignant cells. However, the success of these direct approaches is limited because current technology cannot deliver therapeutic genes to all cancer cells. Alternative strategies have been developed using genes that encode proteins, such as cytokines, which can activate the patient's immune response against the tumour. Another application of gene therapy, which is already undergoing extensive clinical trials, is the transduction of tumour cells with so called suicide genes, which encode enzymes that can convert a prodrug to its toxic metabolite.

Despite the diversity of gene therapies for cancer in both the laboratory and clinic, disappointment is emerging that gene therapy has not fulfilled its early promise.[1] Since no dramatic clinical success has been reported to date, this criticism cannot be ignored, and an objective reconsideration of the potential of gene therapy is timely.

Gene therapy has failed to live up to expectations so far because these have been unrealistic. A lack of realism can be destructive if it leads to disillusionment. Potential benefits of a new approach may be lost if research is abandoned prematurely because dramatic advances have not been made. In oncology, the excitement engendered by gene therapy has been heightened by the inadequacy of current treatments for many of the common adult cancers.

There are good scientific explanations for the inability of gene therapy to produce impressive cures in cancer, but for an understanding of these, gene therapy must be considered within the context and aims of current treatment strategies. Gene therapy provides a prime example of how the medical and scientific communities, working in isolation, can fail to make progress. Doctors may not understand the limitations of the new technology offered by scientists, and scientists may not understand where and how doctors want to apply gene therapy in cancer treatment. By combining scientific and medical approaches, we can define where and why gene therapy is likely to work or fail and explain why gene therapy must not be abandoned before it has had either the time or opportunity to show its worth.

Summary points

- Gene therapy for cancer has not yet fulfilled its early promise
- With current delivery systems, gene therapy is unable to clear large scale disease and is unlikely to become a definitive radical treatment in most clinical scenarios
- Gene therapy will probably be most effective as adjuvant radical treatment through techniques that evoke the "bystander killing effect" and those mediated by the immune system
- Because gene therapy should have fewer side effects than conventional treatments, it may prove effective as a palliative treatment or as a radical combination treatment with radiotherapy or chemotherapy
- Gene therapy should not be abandoned prematurely, appropriate trials within the context of current treatment strategies should still allow it to prove its worth

Current clinical treatments for cancer

When a patient presents with cancer the most important clinical management decision is whether they should receive radical curative or palliative treatment (fig 1). Radical management will comprise a definitive treatment, usually surgery, with additional treatments given before (neoadjuvant) or afterwards (adjuvant). In palliative care the goals are different; toxic treatment is less acceptable and emphasis on relief of symptoms is greater. Scientific evidence suggests that gene therapy can probably contribute to some treatment strategies but is unlikely to make much impact in others.

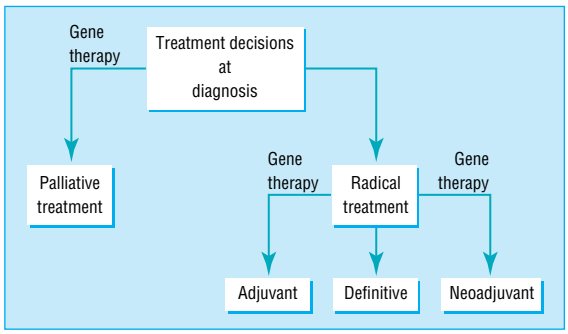

Fig 1 Probable clinical role of gene therapy in palliative and radical treatment for cancer

Radical treatments

Definitive treatments

Patients generally present with a considerable tumour burden in terms of numbers of malignant cells. Animal studies have shown that gene therapy is unlikely to eradicate many tumour cells, even at a single site. This is because current vector systems for delivering genes are inadequate[2] or the immune system, even when activated, is unable to clear large scale disease before it kills its host.[3]

Nevertheless, gene therapy may have a useful role in some tumours, such as high grade glioma, where current treatment is ineffective. In rodent models of glioma, gene therapy was a highly effective radical treatment. Implanted gliomas were eradicated by introducing viral vectors that expressed suicide genes (such as herpes simplex virus thymidine kinase) into the tumour. These instigated tumour cell death by activating prodrugs, such as ganciclovir, which had been given systemically.[4] When these experimental models progressed to human clinical trials, however, the transduction rates of human tumour cells were much lower than those seen in animals.[5]

Adjuvant treatments

Adjuvant treatment targets minimal residual disease and it is here that gene therapy is most likely to be effective as a radical treatment. This is because the number of cancer cells that have to be destroyed is small and gene therapy can overcome the problems it faces in bulk disease.

Optimising the potential of gene delivery systems

The chances of delivering enough copies of a gene to sufficient tumour cells to produce therapeutic effects are greater if the target population is small (see fig 2). With current delivery systems, getting even a single copy of any gene to every tumour cell is unlikely. Because of this we must use gene systems that evoke a "bystander killing effect"—that is, the well documented killing in animal models of non-transduced cells surrounding those tumour cells that have been genetically modified successfully. The bystander effect can be mediated by the passage of toxic metabolites between cells or by the immune system. With help from the bystander effect, the smaller the tumour, the fewer the cells that must be transduced to achieve therapeutic success (fig 2).

Giving the immune system the best chance

In the many gene therapies that seek to activate an immune response (see fig 3), the smaller the tumour the greater the probability that appropriately activated immune cells can control and eradicate it. In animal models of tumour vaccines directed at the immune system, gene therapy was successful only in small tumours.[6]

The problem of immunotherapy in relation to the tumour burden is shown in fig 3. Early in the disease, immunomodulatory gene therapy can eradicate disseminated metastases that are still similar antigenically to the primary tumour and do not yet exert an inhibitory immunosuppressive effect. Immunotherapy is less effective later, as metastases have evolved that are antigenically different from the primary tumour against which the immunity was raised; these suppress the immune response appreciably.

In the clinical setting, specific reasons exist to explain why gene therapy mediated by the immune system is more likely to succeed against small volume disease. In people, tumours evolve over long periods when the immune system is intact. Variants that are poorly immunogenic are probably selected, and these escape clearance by the immune system. The smaller the volume of disease and the shorter the time tumour cells have had to acquire mutations that can make them "invisible," the less likelihood there is that these variants will arise. In addition, various abnormalities of

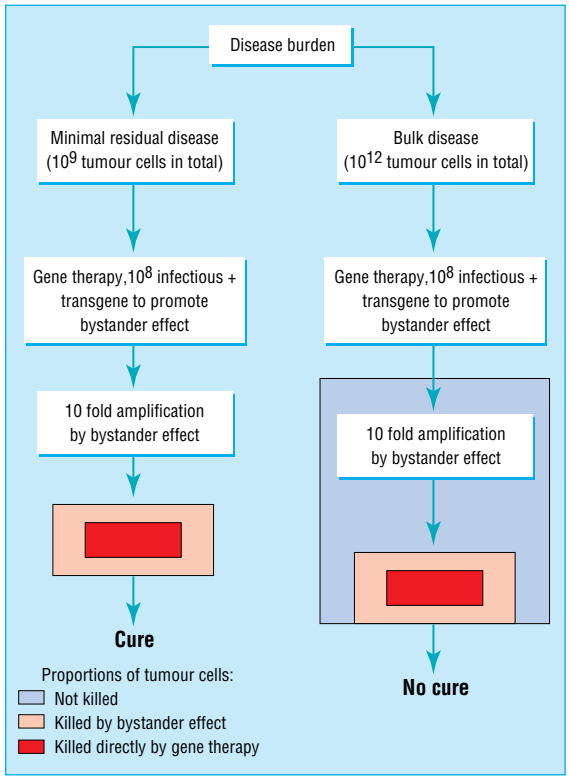

Fig 2 Potential of gene therapy in radical treatment for cancer. In this general scenario, a fixed amount of vector (10^8 infectious units) is delivered in vivo; the vector might encode a transgene which is associated with a bystander effect that allows a 10-fold amplification of the therapeutic benefit of gene delivery, either local or immunologically mediated. In this situation gene therapy is able to cure 10^9 tumour cells (minimal residual disease) but not 10^{12} tumour cells (bulk disease)

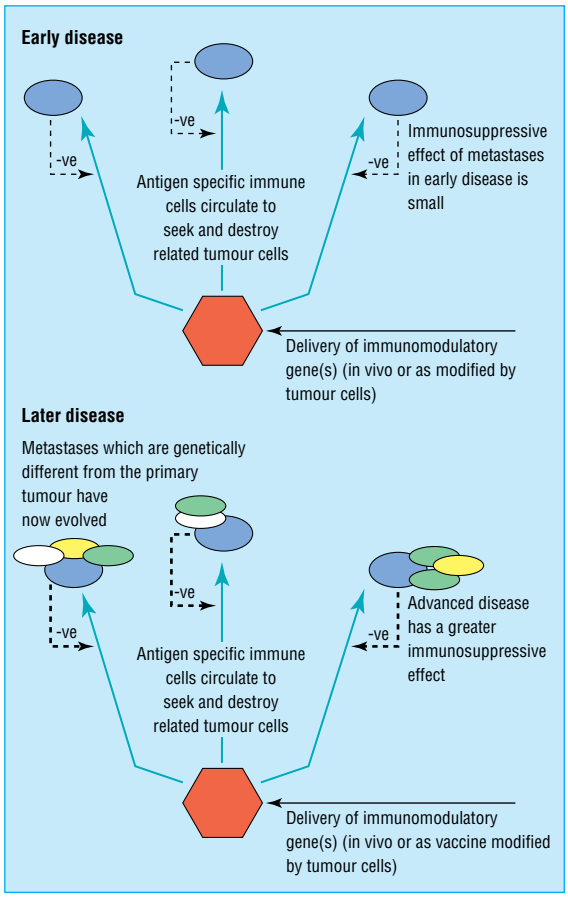

Fig 3 Molecular immunotherapy can be successful in early disease (top) but is problematic in more advanced cancer (bottom)

immune function that correlate with disease extent have been shown in cancer patients, making effective activation of the immune system by gene therapy less likely in advanced disease.[7] Finally, many chemotherapy (and some radiotherapy) regimens suppress the immune system. This compromises further the efficacy of gene therapy in cancers that need extensive conventional treatment.

The way forward in adjuvant treatment

The conclusion that gene therapy is probably effective solely as an adjuvant in radical treatment means that its efficacy can be proved only by large randomised clinical trials with a long follow up. As no such trials are under way, current discontent with curative gene therapy is premature. Nevertheless, future benefit is most likely to be seen in tumours for which there is no current adjuvant treatment after surgery, even in those with a poor prognosis such as deeply invasive melanoma. At least a subset of melanoma in humans seems immunogenic, and this coupled with the poor prognosis is reflected in the large number of phase I and II trials of immunomodulatory gene therapy already started.[8]

Neoadjuvant treatments

Gene therapy may have a role as neoadjuvant treatment. It might be used, for example, to deliver directly into large tumours vectors encoding suicide genes to reduce the number of tumour cells before the definitive treatment. However, the same reservations apply over the efficient delivery of genes to bulk disease and the efficacy of immunotherapeutic gene therapy against large tumours.

Combination treatments

A final option in radical treatment is to deliver two types of treatment simultaneously—for example, giving chemotherapy during radical radiotherapy for head and neck tumours.[9] Established chemotherapy and radiotherapy given in this way can cause considerable morbidity because the patient has side effects from both concurrently. A major theoretical advantage of gene therapy, as discussed below, is that it should have fewer side effects than other treatments. Giving gene therapy concurrently with conventional treatments may therefore be possible. Moreover, since many gene therapy strategies aim to utilise the patients' own immune response against their tumours, the massive antigen release due to death of tumour cells during radiotherapy or chemotherapy may strengthen the effect of immune based gene therapy given simultaneously. This is an area with considerable theoretical benefits, but little experimental work has yet been carried out.

Palliative treatments

Palliative treatments are indicated only if they can relieve symptoms and improve the quality of life. Unfortunately, unpleasant side effects of available treatments often outweigh any small benefits achieved.

One major advantage of gene therapy is that it is specifically directed at tumour cells and treatment toxicity is therefore predicted to be low. Specificity can be derived from a molecular understanding of the inherent genetic defects in a cancer cell[10] and the development of vectors that are transcriptionally,[11] mitotically,[4] or surface targeted.[2] By contrast, radiotherapy and most chemotherapeutic drugs do not specifically target tumour cells. In addition, many molecular immunotherapies seek to recruit the patient's own immune cells as effecters against the tumour and are thus expected to be less toxic than many current standard treatments. Indeed, very few side effects have been found in patients treated with gene therapy to date.[12]

Although gene therapy will be unable to clear all, or even most, of the bulk disease usually found in patients receiving palliative treatment, it may still be effective. Palliative treatments do not need to kill most tumour cells to relieve symptoms. For example, single fraction radiotherapy relieves pain from bone metastases in around 80% of patients, though the percentage of tumour cells killed is probably low.[13]

Palliative gene therapy will probably find a place in specific clinical problems where present treatments fail. It is more likely to be useful, for example, in patients with disseminated disease resistant to chemotherapy such as melanoma or renal cell carcinoma.

The future

From animal models ...

In a few (rare) cases, gene therapies in animal models do give dramatic results. Examples of this include the cure of established rat gliomas by retroviral gene trans-

fer or of melanomas by vaccination with tumour cells secreting cytokine.[14] However, questions arise over the relevance of animal models to cancer patients and the best way of applying these data when designing clinical trials. Results of animal experiments are not usually generalisable to people—mouse tumour cell lines grow more rapidly than human tumours do in vivo, tumours are often not truly syngeneic with their hosts, and the animal's immune system has not been subjected to the prolonged period of selection for poorly immunogenic variants that characterise human cancers.

Despite these reservations, animal models that are relevant to radical adjuvant and palliative treatment in which gene therapy holds most promise have been developed. Treatment can be delivered to an animal shortly after tumour cells have been seeded intravenously but before metastases develop. This mimics the minimal residual disease targeted by adjuvant therapy. Advanced disease typical of that requiring palliative treatment can more easily be modelled. However, gene therapy at this stage, even in animals, is not usually curative.

... to clinical trials

Trying to link laboratory and clinic in these experiments is challenging. Measuring the success of palliative treatment—that is, relief of symptoms—is difficult in animals. Nevertheless, survival curves in animal studies often show a treated group of animals that is not cured of disease but develops tumours appreciably later than the control group. These results are less dramatic (and harder to get published) than those reporting cure, but interpreting them in terms of the treatment's potential in palliation may be useful. Increased life expectancy, with low toxicity and improvement of symptoms, is more than many human cancer treatments can currently achieve. Animal experiments need to be designed, and their data interpreted, with a clearer view of the likely clinical applications in the longer term.

Clinical trials of gene therapies are in the early stages. These are phase I and II studies in patients with advanced, bulky, end stage disease who have already been given a great deal of treatment. All results, including those showing no apparent benefit, must be reported so that ineffective strategies can be abandoned and new approaches devised. If the few positive results reported can be confirmed,[15-17] trials can then be extended to phase III studies in the palliative setting, and from there, it is hoped, to studies of gene therapy as an adjuvant to radical treatment.

Gene therapy should still prove its worth

If gene therapy for cancer has so far failed to live up to expectations, it is because these expectations have been unrealistic. Gene therapy is likely to find a useful place in clinical practice as an adjuvant radical treatment or as a palliative treatment in advanced disease. Unfortunately, it has too often been overpromoted as a radical treatment in advanced disease. Although gene therapy is unlikely to deliver miraculous cures for cancer, it should not be abandoned. Effective progress demands that scientists and clinicians alike must understand each other's capabilities and limitations within the context of defining their common goals. Gene therapy is still in its infancy and although approaches such as vaccination with cytokine transduced, irradiated allogeneic tumour cells expressing known tumour antigens are straightforward and widely applicable, they are perceived as technically demanding and expensive.[18] Genetic intervention in human disease is becoming ever more feasible. In cancer, appropriate preclinical and clinical trials with realistic expectations and designed within the context of present treatment strategies should still allow gene therapy to prove its worth.

Funding: AAM is a Medical Research Council clinical training fellow; IGR is a clinical research fellow supported by the European Society for Medical Oncology. RV and AAM are supported by the Imperial Cancer Research Fund.

Conflict of interest: None.

1 Friedmann T. Human gene therapy—an immature genie, but certainly out of the bottle. *Nature Med* 1996;2:144-7.
2 Zhang J, Russell SJ. Vectors for cancer gene therapy. *Cancer and Metastasis Rev* 1996;15:385-401.
3 Cavallo F, Giovarrelli M, Gulino A, Vacca A, Stoppacciaro A, Modesti A, et al. Role of neutrophils and CD4 + T lymphocytes in the primary and memory response to non immunogenic murine mammary adenocarcinoma made immunogenic by IL-2 gene transfer. *J Immunol* 1992;149:3627-35.
4 Culver KW, Ram Z, Wallbridge S, Ishii H, Oldfield EH, Blaese RM. In vivo gene transfer with retroviral vector-producer cells for treatment of experimental brain tumors. *Science* 1992;256:1550-2.
5 Ram Z. Summary of results and conclusions of the gene therapy of malignant brain tumors: clincal study. *J Neurosurg* 1995;82:343A.
6 Forni G, Cavallo F, Consalvo M, Allione A, Dellabona P, Casorati G, et al. Molecular approaches to cancer immunotherapy. *Cytokines Molecular Ther* 1995;1:225-48.
7 Kavanaugh DY, Carbone DP. Immunologic dysfunction in cancer. *Haematol Oncol Clin N Am* 1996;10:927-51.
8 Carrel S, Johnson JP. Immunologic recognition of malignant melanoma by autologous T lymphocytes. *Curr Opin Oncol* 1993;5:383-9.
9 Dragovic J, Doyle TJ, Tilchen EJ, Nichols RD, Benninger MS, Carlson ER, et al. Accelerated fractionation radiotherapy and concomitant chemotherapy in patients with stage IV inoperable head and neck cancer. *Cancer* 1995;76:1655-61.
10 Mercola D, Cohen JS. Antisense approaches to cancer gene therapy. *Cancer Gene Ther* 1995;2:47-59.
11 Vile RG. Tumour specific gene expression. In: Collins MKL, ed. *Sem Cancer Biol* 1994;5:429-36.
12 Anderson WF. End of the year potpourri—1995. *Hum Gene Ther* 1995;6:1505-6.
13 Allen KL, Johnson TW, Hibbs GG. Effective bone palliation as related to various treatment regimens. *Cancer* 1976;37:984-7.
14 Dranoff G, Jaffee E, Lazenby A, Golumbek P, Levitsky H, Brose K, et al. Vaccination with irradiated tumor cells engineered to secrete murine granulocyte macrophage colony stimulating factor stimulates potent, specific, and long lasting anti-tumor immunity. *Proc Natl Acad Sci USA* 1993;90:3539-43.
15 Marchand M, Weynants P, Rankin E, Arienti F, Belli F, Parmiani G, et al. Tumour regression responses in melanoma patients treated with a peptide encoded by gene MAGE-3. *Int J Cancer* 1995;63:883-5.
16 Izquierdo M, Martin V, de Felipe P, Izquierdo JM, Perez-Higueras A, Cortes ML, et al. Human malignant brain tumour response to herpes simplex thymidine kinase (HSVtk)/ganciclovir gene therapy. *Gene Ther* 1996;3:491-5.
17 Roth JA, Nguyen D, Lawrence DD, Kemp BL, Carrasco CH, Ferson DZ, et al. Retrovirus-mediated wild-type p53 gene transfer to tumors of patients with lung cancer. *Nature Med* 1996;2:985-91.
18 Dalgleish A. The case for therapeutic vaccines. *Melanoma Res* 1996;6:510.

(Accepted 28 May 1997)

Endpiece

Chattering doctors

Doctors should never talk to patients about anything but medicine. When doctors talk politics, economics or sports, they reveal themselves to be ordinary mortals—you know, idiots like the rest of us.

Andy Rooney, quoted in *The Best of Medical Humour* (Howard J Bennett, ed. Philadelphia: Hanley and Belfus, 1997)

Your letter failed to win a place...

Eyal Shahar

Letters that comment on published work are treated differently from the original article itself. They are rarely subject to peer review, and scientific explanations are not usually given when they are rejected. Professor Shahar argues that this is unjustified and counterproductive to scientific inquiry, and that criticism of published work should be subject to peer review.

Division of Epidemiology, University of Minnesota, Minneapolis, MN 55454-1015, USA
Eyal Shahar, *associate professor*
shahar@epivax.epi.umn.edu

BMJ 1997;315:1608–9

The quality of published scientific work is evaluated at least twice—by a handful of reviewers and editors during peer review and by an unknown number of readers after publication. Editorial peer review sometimes helps authors to improve their manuscripts, but more often it helps editors to decide between acceptance and rejection. Just as important—or perhaps even more important—are the unsolicited opinions of readers. Many of the readers are as qualified as the reviewers whose opinions contributed to the editorial decision.[1 2]

The voice of the reader is heard through letters written by the relatively few who formalise their critique in writing. Letters are sometimes better thought out than the original article. They may identify inaccuracies that were missed by formal peer review or uncover flaws in design, analysis, or interpretation. Peer review does not preclude error.[1] Although letter columns are considered important by editorial boards,[3] the fate of correspondence on published work is rarely determined by peer review. The editor(s) usually make the decision whether to publish, and rejection notes to authors are often standardised and contain little, if any, scientific explanation for the decision. Vague statements such as "in the face of fierce competition, your letter failed to win a place" or "many worthwhile contributions must be declined simply for lack of space" are typical.

Anecdotal experience

The example I provide below illustrates the shortcomings of editorial practices. It is a "letter to the editor" (coauthored by Paul G McGovern) that was rejected with no specific explanation. Since the letter challenges unequivocally the main conclusion of an article, either the challenge or the original conclusion must have been faulty. The letter can, therefore, only be rejected on the assumption that its contents are faulty. I am asking the reader to judge this assertion (ignoring the question of whether the letter's content is true or false). I also provide the letter as a test case of my proposal for peer review of correspondence, inviting concrete evaluation of its content that will either support or oppose the editorial verdict in this case.

Discussion

Only two legitimate reasons exist for the journal to reject this letter—most of its content was judged faulty or the editor(s) preferred to publish another letter with a similar message. Lack of space should not justify rejection because space should be made available for corrections, even at the expense of delaying the publication of new original articles. I saw no letter with a similar message in follow up correspondence.[5 6]

When a manuscript is rejected by a journal, the authors may get it published elsewhere. When a letter

Letter to the editor

Sir—Andreotti et al conclude that "among patients with acute myocardial infarction, those with prodromal unstable angina ... have remarkably faster responses to treatment with tissue plasminogen activator than those without such symptoms."[4] We take issue with their inference.

The study by Andreotti et al was an observational cohort study of 23 patients who had suffered a myocardial infarction. Of these patients, fourteen had experienced preinfarction angina and 9 had not ("exposed" and "unexposed" groups, respectively, in epidemiological terminology.) The only acceptable inference from this small study is that exposure to unstable angina before a myocardial infarction was associated with more rapid reperfusion and smaller infarcts than were observed in the absence of prodromal unstable angina. That the two groups of patients happened to be treated with tissue plasminogen activator (and by other means) further defines the cohort characteristics, but is irrelevant to the question of whether thrombolytic therapy is more effective in the presence of preinfarction angina than in its absence. As the authors acknowledge, reperfusion occurs spontaneously during the course of myocardial infarction and therefore its rate is far from being entirely determined by treatment with tissue plasminogen activator. For example, Andreotti et al would have observed exactly the same results if treatment with tissue plasminogen activator had identical effects in both groups of patients yet infarctions that follow unstable angina are associated with faster rates of spontaneous reperfusion (and tend to be smaller) than infarctions without prodromal unstable angina.

To show a differential effect of treatment with tissue plasminogen activator in the presence (versus absence) of preinfarction angina, one should demonstrate a statistically significant interaction between two effects: the effect of treatment with tissue plasminogen activator in myocardial infarction patients who had preinfarction angina and the effect of treatment with tissue plasminogen activator in those who did not. Each of these effects can only be estimated by comparing the reperfusion rate and infarction size in patients treated with tissue plasminogen activator to these measures in patients who were not, preferably by a randomised design. Unfortunately, such a design is no longer feasible since it is ethically unacceptable to withhold thrombolytic therapy from patients who should receive it.

criticising a published article is rejected no such remedy is usually available. In this sense, an erroneous editorial decision to reject a letter may be more damaging to scientific progress than an erroneous decision to reject a manuscript.

The differential treatment of scientific correspondence and manuscripts is not unique to scientific journals. Scientists rarely cite criticism of original research,[2] and academic institutions give little or no credit for published letters.

The most truthful message in any particular case cannot be deduced from some general rule of importance, even if there were an empirical way of substantiating such a rule. What is important is not the origin of the message (for example, authors of a manuscript or authors of a letter) but the message itself. Is it scientific or perhaps pseudoscientific? Does it survive logical criticism or not? Differential treatment of scientific communications introduces a potential prejudice into the search for objective knowledge.

Peer review of manuscripts is based on criteria such as clarity, validity, originality, and relevance. Peer review of correspondence could follow the same path. Letters to the editor (including the reply of the authors of the original article, which usually escapes rejection) should be evaluated for their scientific merit, and their fate should be determined on specific grounds. A letter may be rejected, for example, because its argument is judged to be rhetorical, its content faulty, or the thoughts of the author poorly articulated. Unexplained decisions leave too much room for speculation and, sometimes, suspicion.

Opponents of my suggestion for peer review of correspondence may argue that the process is lengthy and that it is essential to publish follow up correspondence quickly, while the original article is still fresh in the reader's memory. Scientific progress, however, is not a race against an arbitrary deadline.

It might be argued that formulating specific criteria for evaluating letters might be difficult. But a scientific communication—that is, one that claims to advance objective knowledge—should lend itself to critical appraisal, above and beyond just "feeling" for its merit.

Some might claim that peer review for letters could lead to an infinite, regressive process of publishing letters that comment on letters, and that such a process would have to be stopped arbitrarily anyway. Most debate, however, fades naturally away after one or two rounds, and if it does not, peer review should be responsible for identifying reiterative stages of a correspondence and for stopping it. Remember too that arbitrary termination also happens with peer review of manuscripts since the reviewers' critiques are not subject to peer review.

In a recent article, Bhopal and Tonks asked, "If published critical comment is considered integral to research should it not be peer reviewed?"[2] My answer is "Yes, it should." Editors would do justice to science if they solicited peer review of correspondence, including peer review of the reply from the authors of the original article. Editors who object should provide authors of rejected letters to the editor with their own scientific review to support their decision.

"Errors may lurk even in our best tested theories. It is the responsibility of the professional to search for these errors."
Neil McIntyre, Karl Popper

I thank Lori Vitelli, Maureen Smith, Jacqueline Dekker, the anonymous reviewer, and the editorial board of the BMJ for helpful comments on earlier versions of the article.

1 Spodick DH. The peer review system and the editor's correspondence. *Arch Intern Med* 1981;141:1121.
2 Bhopal RS, Tonks A. The role of letters in reviewing research: always look for letters that follow original papers. *BMJ* 1994;308:1582-3.
3 International Committee of Medical Journal Editors. Statements from the Vancouver Group. *BMJ* 1989;299:1394-5.
4 Andreotti F, Pasceri V, Hackett DR, Davies GJ, Haider AW, Maseri A. Preinfarction angina as a predictor of more rapid coronary thrombolysis in patients with acute myocardial infarction. *N Engl J Med* 1996;334:7-12.
5 Migrino RQ, Moliterno DJ, Topol EJ. Preinfarction angina. *N Engl J Med* 1996;335:59.
6 Kloner RA, Gibson M, Cannon C, Braunwald E. Preinfarction angina. *N Engl J Med* 1996;335:59-60.

(Accepted 4 March 1997)

The editor's decision is final

Liz Crossan, Richard Smith

Last year the *BMJ* received about 3850 letters for publication. We had space to publish about 1200 in the paper journal. We cannot increase the number of pages we devote to letters for two reasons—readers do not want us to and we could not afford to do so. We give priority to letters offering cogent criticisms of material we have already published, but we cannot publish all these letters. Some have to be rejected.

We hope soon to post almost all letters on our web site within days of receiving them. Only those that are libellous, obscene, or incomprehensible will be excluded. We will then make a selection in the normal way for publication in the paper journal, aided perhaps by comments on the letters posted on the web site.

We do not plan to peer review externally all letters as Dr Shahar suggests, mainly because we see letters as a form of peer review. As Dr Shahar points out, we do not externally review the comments of peer reviewers. We have to stop somewhere, and we choose to stop with peer reviewers' comments and with letters for publication in response to published material. Again this is a question of resources—peer reviewing all letters externally would be time consuming and expensive.

Ideally, we would offer all authors of rejected papers and letters a specific explanation for the rejection. Again, we don't do this because of resources. We do not want to increase the price of the journal (and if we did our resources might be reduced because fewer people might subscribe), and we want to concentrate resources on improving the quality of what we do publish rather than on supplying justification for rejecting material. We will always, however, explain why something has been rejected when authors ask us to do so.

British Medical Journal, London WC1H 9JR
Liz Crossan, *letters editor*
Richard Smith, *editor*
Correspondence to: Ms Crossan

Meta-analysis

Beyond the grand mean?

George Davey Smith, Matthias Egger, Andrew N Phillips

This is the third in a series of six articles examining the procedures in conducting reliable meta-analysis in medical research

Department of Social Medicine, University of Bristol, Bristol BS8 2PR
George Davey Smith, *professor of clinical epidemiology*
Matthias Egger, *reader in social medicine and epidemiology*
Department of Primary Care and Population Sciences, Royal Free Hospital School of Medicine, London NW3 2PF
Andrew N Phillips, *professor of epidemiology and biostatistics*
Correspondence to: Professor Davey Smith zetkin@bristol.ac.uk

BMJ 1997;315:1610–4

In the previous two articles[1,2] we outlined the potentials and principles of meta-analysis and the practical steps in performing a meta-analysis. Now we will examine how to use meta-analysis to do more than simply combine the results from all the individual trials into a single effect estimate. Firstly, we discuss the advantages and disadvantages of performing subgroup analyses. Secondly, we consider the situation in which the differences in effects between individual trials are related in a graded way to an underlying phenomenon, such as the degree of mortality risk of the trial participants.

Summary points

- Meta-analysis can be used to examine differences in treatment effects across trials; however, the fact that randomised trials are included in meta-analyses does not mean that comparisons between trials are also randomised comparisons
- Meta-analytic subgroup analyses, like subgroup analyses within trials, are prone to bias and need to be interpreted with caution
- A more reliable way of assessing differences in treatment effects is to relate outcome to some underlying patient characteristic on a continuous, or ordered, scale
- The underlying level of risk is a key variable which is often related to a given treatment effect, with patients at higher risk receiving more benefit then low risk patients
- Individual patient data, rather than published summary statistics, are often required for meaningful subgroup analyses

Subgroup analysis

The main aim of a meta-analysis is to produce an estimate of the average effect seen in trials of a particular treatment. The direction and magnitude of this average effect is intended to guide decisions about clinical practice for a wide range of patients. Clinicians are thus being asked to treat their patients as though each one is well represented by the patients in the clinical trials included in the meta-analysis. This runs against doctors' concerns to use the specific characteristics of a patient to tailor that patient's management.[3] Indeed, the effect of a given treatment is unlikely to be identical across different groups of patients—for example, young people versus elderly people, those with mild disease versus those with severe disease. It may therefore seem reasonable to base treatment decisions on the results of the trials that have included participants with similar characteristics to the patient under consideration rather than on the overall evidence as provided by meta-analysis.

Decisions based on subgroup analyses, however, are often misleading. Consider, for example, a doctor in Germany being confronted by the meta-analysis of long term β blockade after myocardial infarction (see previous article[2]). Although a robust beneficial effect is seen in the overall analysis, in the only trial that recruited a substantial proportion of German patients (trial N in previous article),[4] there was, if anything, a detrimental effect associated with β blockers. Should the doctor give β blockers to German patients who have had an infarction? Common sense suggests that being German does not prevent a patient from obtaining benefit from β blockade. Thus the best estimate of the outcome for German patients may come through discounting the trial carried out in German patients. This may seem paradoxical; indeed the statistical expression of this phenomenon is known as Stein's paradox (box).[5]

Making decisions between overall effects and particular results is not just a problem created by meta-analysis; it also applies to the interpretation of individual clinical trials.[6] Authors of trial reports often spend more time discussing the results seen in subgroups of patients included in the trial than on the overall results. Yet frequently the findings of these subgroup analyses fail to be confirmed by later research. The various trials of β blockade after myocardial infarction yielded several subgroup findings with apparent clinical significance.[7] Treatment was said to be beneficial in patients aged under 65 but harmful in older patients, or only beneficial in patients with anterior myocardial infarction. When examined in subsequent studies or in a formal pooling project[8] these findings received no support.[7] It can be shown that if an overall treatment effect is significant at the 5% level ($P < 0.05$) and the patients are divided at random into two similarly sized groups then there is a 1 in 3 chance that the treatment effect will be large and highly significant in one group but irrelevant and non-significant in the other.[9] Which subgroup "clearly" benefits from an intervention is thus often a chance phenomenon, inundating the literature with contradictory findings from subgroup analyses and wrongly inducing clinicians to withhold treatments from some patients.[10-12]

Meta-analyses offer a sounder basis for subgroup analysis, but they are not exempt from producing misleading findings. One of the explanations for the disappointing result seen in the β blocker trial in German patients was that the agent used, oxprenolol, had intrinsic sympathomimetic activity.[13] This seemed plausible because the beneficial effect was assumed to be entirely mediated by blockade of the β 1 receptor, and the supposition was supported by subgroup analysis in a meta-analysis,[14,15] which showed less benefit in trials of patients treated with agents with intrinsic sympatho-

Stein's paradox

- Applying the findings from meta-analyses often means that the results from a particular trial are disregarded in favour of the combined result. This will generally be based on the assumption that inconsistent results are purely due to chance. But even if some real differences exist the overall estimate may still provide the best estimate of the effect in that group (Stein's paradox).[5]
- Charles Stein showed that a quantity can be better estimated by taking into account the findings from similar studies, rather than by basing estimation solely on one study. The central principle of Stein's method is the "shrinking" of individual data points towards the grand mean. The amount by which an observed value is adjusted (shrinking factor) will depend on the precision of this value. An outlying value that was measured imprecisely is shrunk towards the grand mean to a greater extent than an outlier that was measured with considerable precision. The result of the trial including German patients contributed only little weight in the combined analysis (see the main text)[4] and would thus be shrunk a long way towards the overall estimate of a beneficial effect of β blockade.

mimetic activity (fig 1). The difference between the two classes of β blockers was significant ($P<0.01$). Since then, however, a trial was published showing a particularly strong beneficial effect of acebutolol, an agent with intrinsic sympathomimetic activity,[16] whereas another trial using metoprolol, a β blocker without intrinsic sympathomimetic activity, was essentially negative.[17] This illustrates that, far from aiding clinicians, post hoc subgroup analyses may confuse and mislead. A more reliable way of assessing differences in treatment effects is to relate outcome to some underlying patient characteristic on a continuous, or ordered, scale.[18 19]

Meta-regression: examining gradients in treatment effects

The clinical trials included in a meta-analysis often differ in a way that would be expected to modify the outcome. In trials of cholesterol reduction the degree of cholesterol lowering attained differs markedly between studies, and the reduction in mortality from coronary heart disease is greater in the trials in which larger reductions in cholesterol are achieved.[18 20] Such graded associations are not limited to situations where greater benefits would be expected consequent on greater changes in a risk factor. In the case of thrombolysis after acute myocardial infarction, the greater the delay in treatment, the smaller the benefit of thrombolysis.[21 22] Here, the graded association is seen between the outcome and a characteristic of the treatment used. Such a gradient allows for a more powerful examination of differences in outcomes, as a statistical test for trend can be performed, rather than the less powerful test for evidence of global heterogeneity. Other attributes of study groups—such as age and length of follow up—can readily be analysed in this way. As discussed later in this series,[23] such analyses will often require data on individual patients rather than published summary statistics.

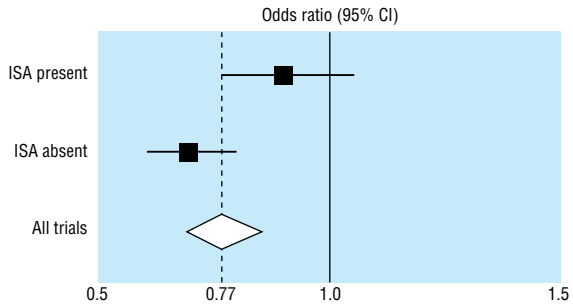

Fig 1 Total mortality from trials of β blockers in secondary prevention after myocardial infarction. Meta-analysis stratified by presence or absence of intrinsic sympathomimetic activity (ISA). Adapted from Yusuf et al[14]

Risk stratification

A factor that is often related to a given treatment effect is the underlying risk of occurrence of the event that the treatment aims to prevent. It makes intuitive sense that patients at high risk are more likely to benefit than those at low risk. In the case of trials of cholesterol lowering, for example, the patient groups have ranged from survivors of heart attack with gross hypercholesterolaemia to groups of healthy asymptomatic people with moderately raised cholesterol concentrations. The death rates from coronary heart disease in the first group have been up to 100 times higher than the death rates in the second groups. The outcome of treatment in terms of all cause mortality has been more favourable in the trials recruiting participants at high risk than in the trials recruiting participants at relatively low risk.[18] Two factors contribute to this. Firstly, among the high risk participants, the great majority of deaths will be from coronary heart disease, the risk of which is reduced by cholesterol reduction. A 30% reduction in mortality from coronary heart disease therefore translates into a near equivalent reduction in total mortality. In the low risk participants, on the other hand, a much smaller proportion—about 40%—of deaths will be from coronary heart disease. In this case a 30% reduction in mortality from coronary heart disease would translate into a much smaller—about 10%—reduction in all cause mortality. Secondly, if there is any detrimental effect of treatment it may easily outweigh the benefits of cholesterol reduction in the low risk group, whereas in high risk patients, among whom a substantial benefit is achieved from cholesterol reduction, this will not be the case. In a recent meta-analysis of cholesterol lowering trials this situation was evident for trials using fibrates but not for trials using other drugs.[24]

A similar association between level of risk and benefit can be seen in meta-analyses carried out for other types of medical treatment.[25] Thus the use of antiplatelet agents such as aspirin produces a 23% reduction in all cause mortality after an acute myocardial infarction but only a (non-significant) 5% reduction in the primary prevention setting.[26] This may reflect a small increase in the risk of haemorrhagic stroke consequent on the use of antiplatelet agents, which counterbalances the beneficial effects on coronary heart disease among low risk individuals but not among those at higher risk. Similarly, a large reduction in relative risk

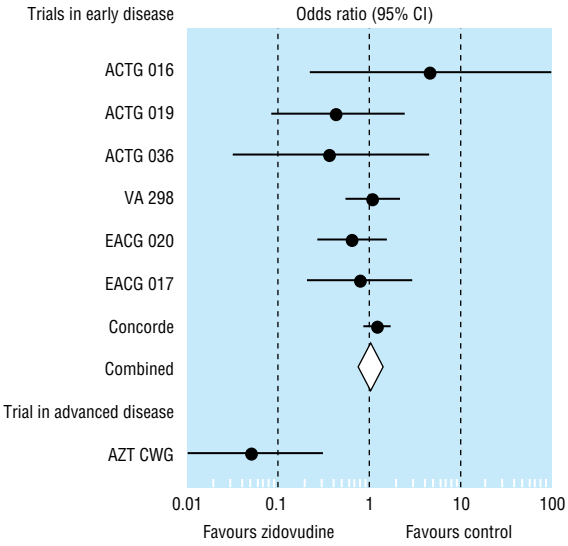

Fig 2 Meta-analysis of mortality results of trials of zidovudine in asymptomatic or early symptomatic HIV infection. The results are in stark contrast to the beneficial effect seen in the only trial in high risk patients (AZT Collaborative Working Group).[27] Adapted from Egger et al[28]

of death was seen in the single study that has reported on treating HIV infection with zidovudine in patients with AIDS.[27] A meta-analysis of seven trials, however, showed that zidovudine given early in the course of HIV infection was not associated with any long term survival benefit (fig 2).[28] When outcomes are very different in groups at different levels of risk it is inappropriate to perform a meta-analysis in which an overall estimate of the effect of treatment is calculated. In the zidovudine trials, for example, an overall effect estimate from all eight trials (odds ratio 0.96; 95% confidence interval 0.75 to 1.22) is very different from that seen in the only trial among patients with AIDS (0.04; 0.01 to 0.33). If there had been more trials among patients with AIDS the overall effect would seem highly beneficial. Conversely, if there had been more large trials among asymptomatic patients the confidence limits around the overall effect estimate would exclude any useful benefit, which would be misleading if applied to patients with AIDS.

Problems in risk stratification

When many trials have been conducted in a particular field, risk stratification can be performed at the level of individual trials. This was carried out in the case of cholesterol lowering, with mortality from coronary heart disease in the control arm of the trials as the stratification variable.[18] This stratification is of clinical use, as this is the risk of death from coronary heart disease in patients without treatment—that is, the risk level that clinicians want to use for deciding whether patients will benefit from therapeutic cholesterol lowering. The analysis can also use risk of death in the control group as a continuous variable, through the examination of the interaction between treatment effect and risk in a logistic regression analysis. A significant statistical test for interaction suggests that there is a real difference in outcome at different levels of risk.

The use of mortality in the control group as a stratification variable introduces a potential bias into the analysis, as this mortality is included in the calculation of the effect estimate from each trial.[18 29-31] Thus, if through chance variation, mortality from coronary heart disease in the control group happens to be low, apparently unfavourable effects of the treatment on mortality would be likely, as mortality in the treatment group would apparently be increased. This would itself produce an association between the outcome measure and the level of risk in the control group, with greater benefit (and fewer disbenefits) being seen in the trials in which the play of chance led to a high mortality in the control group. For example, a recent meta-regression analysis examined whether in middle aged patients with mild to moderate hypertension the benefit from drug treatment depends on the underlying risk of death.[32] The scatterplot advocated by L'Abbé et al[33] of event rates in the treated group against those in the control group was used (fig 3 (top)). This plot is useful for examining the degree of heterogeneity between trials and to identify outliers. If the treatment is beneficial, trials will fall to the right of the line of identity (the no effect line). A homogenous set of trials will scatter around a parallel line, which corresponds to the combined treatment effect.

The authors then computed a linear regression model describing mortality in the treated groups as a function of mortality in the control group.[32] Because the number of deaths and person years of follow up varied widely between studies, the analysis was weighted by the inverse of the variance of the rate ratio. The resulting regression line intersects with the "null effect" line at a rate of 6 per 1000 person years in the control group (fig 3 (top)). This was interpreted as showing "that drug treatment for mild to moderate hypertension has no effect on, or may even increase, all cause mortality in middle aged patients." [32] In other words, antihypertensive treatment was considered to be beneficial only in patients at relatively high risk of death. This interpretation, however, is misleading because it ignores the influence of random fluctuations on the slope of the regression line.[29] If, owing to non-infinite sample sizes, mortality in a control group is particularly low then mortality in the treatment group will, on average, seem high. Conversely, if mortality among controls is by chance high then mortality in the treatment group will seem low. The effect of random error will thus rotate the regression line around a pivot, making it cross the line of identity on the right hand side of the origin.

This phenomenon, a manifestation of regression to the mean,[30] can be illustrated in computer simulations. Using the same rates in the control group and assuming a constant reduction of all cause mortality of 10% in treated groups (relative risk 0.9), we considered the situation both assuming no random fluctuations in rates and allowing random error (fig 3 (bottom)).[29] After we added error (by sampling 1000 times from the corresponding Poisson distribution) the regression line rotated and crossed the no effect line. Indeed, the intersection is at almost the same point as that found in the earlier meta-analysis—namely, at a mortality in the control group of about 6 per 1000 person years. It is thus quite possible that what was interpreted as reflecting detrimental effects of antihypertensive treatment[32]

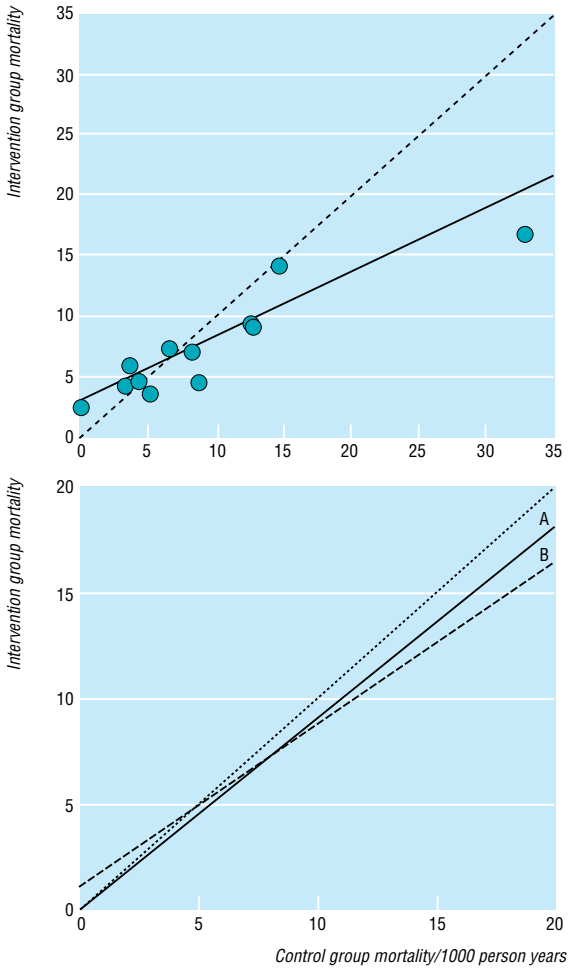

Fig 3 Top: All cause mortality in intervention and control groups of clinical trials in mild to moderate hypertension.[32] The dotted line represents no effect, with identical mortality in both groups, and the solid line represents the weighted regression line. Bottom: Computer simulation based on the same trials. Line A assumes a constant relative risk reduction of 10%. Line B corresponds to Line A after random error was added to the mortality. Adapted from Egger et al[29]

was in fact produced by random variation in event rates.

When mortality in the control groups vary greatly or when trials are large, the chance fluctuations that produce such spurious associations will be less important. Alternatively, the analysis can be performed using the overall mortality in the control and treatment arms of the trials as the risk indicator.[18] This will generally, but not always, lead to bias in the opposite direction, diluting any real association between level of risk and treatment effect.[30]

Use of event rates from either the control group or overall trial participants as the stratifying variable when relating treatment effect to level of risk is thus problematic.[29 30] Although some, more complex, statistical methods are less susceptible to these biases,[31 34] it is preferable to use indicators of risk that are not based on outcome measures. In the case of the effect of angiotensin converting enzyme inhibitors on mortality in patients with heart failure, use of risk in the control group showed greater relative and absolute benefit in trials recruiting higher risk participants.[25] In a meta-analysis, data were available on treatment effects according to clinical indicators within strata from many of the trials.[35] Twenty nine per cent of patients with an ejection fraction of ≤ 0.25 at entry died during the trials, compared with 17% of patients with an ejection fraction of >0.25. A substantial reduction in mortality (odds ratio 0.69; 95% confidence interval 0.57 to 0.85) was seen in the first, higher risk group, whereas little effect on mortality was seen in the second, lower risk group (0.98; 0.79 to 1.23). A similar difference was seen if a combined end point of mortality or admission to hospital for congestive heart failure was used as the outcome measure.

Confounding

That randomised controlled trials are included in meta-analyses does not mean that comparisons made between trials are randomised comparisons. When outcomes are related to characteristics of the trial participants, to differences in treatments used in the separate trials, or to the situations in which treatments were given, the associations seen are subject to the potential biases of observational studies. Confounding could exist between one trial characteristic—say, drug trials versus diet trials in the case of cholesterol lowering—and another characteristic, such as level of risk of the participants in the trial. In many cases there are simply too few trials, or differences in the average characteristics of participants in the trials are too small, for a stratified analysis to be performed at the level of the individual trial. It may be possible to consider strata within the trials—for example, male versus female, or those with or without existing disease—to increase the number of observations to be included in the regression analysis. Increasing the number of data points in this way is of little help if there are strong associations between the factors under consideration. For example, in a meta-regression analysis of total mortality outcomes of cholesterol lowering trials various factors seem to influence the outcome: greater cholesterol reduction leads to greater benefit; trials including participants with a higher level of risk of coronary heart disease show larger mortality reductions; and the fibrate drugs lead to less benefit than other interventions.[20 24] These findings are difficult to interpret, however, as the variables included are strongly related—fibrates have been used mainly in trials recruiting lower risk participants, and they lower cholesterol much less than statins. In this situation all the problems of performing multivariable analyses with correlated covariates are introduced.[36 37]

Conclusion

It is tempting to use a meta-analysis to produce more than a simple overall effect estimate, but caution is needed, for the reasons detailed above. One of the more useful extensions of meta-analysis beyond the grand mean relates to the examination of publication bias and other inclusion biases, which will be discussed later in this series.

The department of social medicine at the University of Bristol and the department of primary care and population sciences at the Royal Free Hospital School of Medicine, London, are part of the Medical Research Council's health services research collaboration.

Funding: ME was supported by the Swiss National Science Foundation.

1 Egger M, Davey Smith G. Meta-analysis: potentials and promise. *BMJ* 1997;315:1371-4.
2 Egger M, Davey Smith G, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997;315:1533-7.
3 Wittes RE. Problems in the medical interpretation of overviews. *Stat Med* 1987;6:269-76.
4 The European Infarction Study Group. European infarction study (EIS). A secondary prevention study with slow release oxprenolol after myocardial infarction: morbidiy and mortality. *Eur Heart J* 1984;5:189-202.
5 Efron B, Morris C. Stein's paradox in statistics. *Sci Am* 1977;236:119-27.
6 Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992;116:78-84.
7 Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991;266:93-8.
8 Beta-Blocker Pooling Project Research Group. The beta-blocker pooling project (BBPP): subgroup findings from randomized trials in post infarction patients. *Eur Heart J* 1988;9:8-16.
9 Peto R. Statistical aspects of cancer trials. In: Halnan KE, ed. *Treatment of cancer*. London: Chapman and Hall, 1982.
10 Buyse ME. Analysis of clinical trial outcomes: some comments on subgroup analyses. *Controlled Clinical Trials* 1989;10:187-94S.
11 Mauri F, Gasparini M, Barbonaglia L, Santoro E, Franzosi M, Tognoni G, et al. Prognostic significance of the extent of myocardial injury in acute myocardial infarction treated by streptokinase (the GISSI trial). *Am J Cardiol* 1989;63:1291-5.
12 Peto R. Misleading subgroup analysis in GISSI. *Am J Cardiol* 1990;64:771
13 Schroeder R. Oxprenolol in myocardial infarction survivors: brief review of the European infarction study results in the light of other beta-blocker post infarction trials. *Z Kardiol* 1985;74(suppl 6):165-72.
14 Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Progr Cardiovasc Dis* 1985;17:335-71.
15 Peto R. Why do we need systematic overviews of randomized trials? *Stat Med* 1987;6:233-40.
16 Boissel JP, Leizorovicz A, Picolet H, Peyrieux JC. Secondary prevention after high-risk acute myocardial infarction with low-dose acebutolol. *Am J Cardiol* 1990;66:251-60.
17 Lopressor Intervention Trial Research Group. The lopressor intervention trial: multicentre study of metoprolol in survivors of acute myocardial infarction. *Eur Heart J* 1987;8:1056-64.
18 Davey Smith G, Song F, Sheldon TA. Cholesterol lowering and mortality: the importance of considering initial level of risk. *BMJ* 1993;306:1367-73.
19 Bailey KR. Generalizing the results of randomized clinical trials. *Controlled Clinical Trials* 1994;15:15-23.
20 Holme I. Relationship between total mortality and cholesterol reduction as found by meta-regression analysis of randomized cholesterol lowering trials. *Controlled Clinical Trials* 1996;17:13-22.
21 Fibrinolytic Therapy Trialists' (FTT) Collaborative Group. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients. *Lancet* 1994;343: 311-22.
22 Zelen M. Intravenous streptokinase for acute myocardial infarction. *New Engl J Med* 1983;308:593.
23 Davey Smith G, Egger M. Meta-analysis: unresolved issues and future developments. *BMJ* (in press).
24 Davey Smith G. Low blood cholesterol and non-atherosclerotic disease mortality: where do we stand? *Eur Heart J* 1997;18:6-9.
25 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.
26 Antiplatelet Trialist's Collaboration. Collaborative overview of randomised trials of antiplatelet therapy—I: prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. *BMJ* 1994;308:81-106.
27 Fischl MA, Richman DD, Griego MH, Gottlieb MS, Volberding PA, Laskin OL, et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *N Engl J Med* 1987;317:185-91.
28 Egger M, Neaton JD, Phillips AN, Davey Smith G. Concorde trial of immediate versus deferred zidovudine. *Lancet* 1994;343:1355.
29 Egger M, Davey Smith G. Risks and benefits of treating mild hypertension: a misleading meta-analysis? *J Hypertens* 1995;13:813-5.
30 Sharp SJ, Thompson SG, Altman DG. The relation between treatment benefit and underlying risk in meta-analysis. *BMJ* 1996;313:735-8.
31 Thompson SG, Smith TEC, Sharp SJ. Investigating underlying risk as a source of heterogeneity and meta-analysis. *Stat Med* (in press).
32 Hoes AW, Grobbee DE, Lubsen J. Does drug treatment improve survival? Reconciling the trials in mild-to-moderate hypertension. *J Hypertens* 1995;13:805-11.
33 L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;107:224-33.
34 McIntosh MW. The population risk as an explanatory variable in research synthesis of clinical trials. *Stat Med* 1996;15:1713-28.
35 Garg R, Yusuf S, for the Collaborative Group on ACE Inhibitor Trials. Overview of randomised trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. *JAMA* 1995;273:1450-6.
36 Phillips AN, Davey Smith G. How independent are "independent" effects? Relative risk estimation when correlated exposures are measured imprecisely. *J Clin Epidemiol* 1991;44:1223-31.
37 Davey Smith G, Phillips AN. Confounding in epidemiological studies: why "independent" effects may not be all they seem. *BMJ* 1992;305: 757-9.

BMJ audit: time to decision and publication

We aim to make a decision on publication within eight weeks (56 days); to reject papers that are unsuitable for external peer review within two weeks (14 days); and to publish a paper within eight weeks of acceptance.

Between July and December last year we made a decision within 56 days for 68% of all papers submitted (1462/2141) and for 35% of those accepted (93/269). We accepted 45% within 66 days, and the mean time to accept a paper was 81 days. We met our target of rejecting papers without peer review within 14 days for 41% of papers (495/1207); 72% were rejected within 24 days.

Overall we published 24% of papers within eight weeks of acceptance, 42% within 10 weeks, and 61% within 12 weeks. Of research papers, however, we published 21% within eight weeks, 51% within 10 weeks, and 79% within 12 weeks.

Table 1 Results of *BMJ* audits. Values are percentages unless stated otherwise

	Decision within 56 days		Accepted papers		Rejected papers (no peer review)		Publication after acceptance within:		
Audit	All papers	Accepted papers	Decision within 66 days	Mean time to decision (days)	Decision within 24 days	Mean time to decision (days)	8 weeks	10 weeks	12 weeks
1993:									
Jan-June	88	73	85	41	76	19	38	75	95
July-Dec	86	62	75	50	84	18	27	66	85
1994:									
Jan-June	88	64	76	48	84	18	13	24	57
July-Dec	83	64	73	51	73	21	40	67	87
1995:									
Jan-June	72	41	53	69	56	26	38	60	76
July-Dec	73	34	43	81	65	22	32	50	73
1996:									
Jan-June	81	43	59	59	65	24	19	35	53
July-Dec	76	53	56	88	62	23	29	52	66
1997:									
Jan-June	68	35	45	81	72	20	24	42	61