who denied making changes in the bedroom. This raises the intriguing possibility that non-feather substitutes may pose a greater risk of asthma than any allergens associated with feather bedding.

The single "bedroom alterations" question (on the 1993 questionnaire) may have been insufficient to detect all conscious action by families to ensure that their children do not have feather pillows. Nevertheless, to explain the inverse association of asthma with feather pillows purely on the basis of undetected avoidance it would be necessary to assume that more than half of the families of cases who denied bedroom alterations and who would otherwise have used feather pillows avoided doing so because of their child's chest trouble (table II). This seems unlikely. Further epidemiological studies including more specific inquiries about allergen avoidance are required to confirm the apparent risk associated with non-feather pillows. If this association is causal, however, we estimate that it accounts for 53% of the severe asthma in our population.

Although firm conclusions cannot be drawn from a single study, our results suggest that avoidance of feather bedding is unlikely to benefit children with asthma and raise the possibility that there is a hitherto unidentified hazard associated with synthetic pillows. Volatile organic compounds released in low concentrations close to the breathing zone might increase mucosal permeability to inhaled allergens and thereby offer a speculative explanation for a possible causal link between foam pillows and childhood asthma. This possibility should be investigated by further observational and experimental studies.

We thank all the school staff and parents who took part in the 1991 and 1993 surveys; Dr Elizabeth Taylor and Ms Jayne Thompson for assistance with the 1991 survey; and Dr Jane

Scarlett and Ms Joanne Griffiths for help with the 1993 mailing.

Funding: The National Asthma Campaign.

Conflict of interest: None.

- 1 Wardlaw AJ. The role of air pollution in asthma. *J Allergy Clin Immunol* 1992;90:358-63.
- 2 Infante-Rivard C. Childhood asthma and indoor environmental risk factors. *Am J Epidemiol* 1993;137:834-44.
- 3 Sears MR, Herbison GP, Holdaway MD, Hewitt CJ, Flannery EM, Silva PA. The relative risks of sensitivity to grass pollen, house dust mite and cat dander in the development of childhood asthma. *Clin Exp Allergy* 1989;19:419-24.
- 4 Strachan DP, Anderson HR, Limb ES, O'Neill A, Wells N. A national survey of asthma prevalence, severity, and treatment in Great Britain. *Arch Dis Child* 1994;70:174-8.
- 5 Brunekreef B, Groot B, Hoek G. Pets, allergy and respiratory symptoms in children. *Int J Epidemiol* 1992;21:338-42.
- 6 Pearce N, Weiland S, Keil U, Langridge P, Anderson HR, Strachan DP, et al. Self reported prevalence of asthma symptoms in Australia, England, Germany and New Zealand: an international comparison using the ISAAC protocol. *Eur Respir J* 1993;6:1455-61.
- 7 Francis B, Green M, Payne C, eds. *The GLIM system release 4 manual*. Oxford: Clarendon Press, 1993.
- 8 Linna O. Environmental and social influences on skin test results in children. *Allergy* 1983;38:513-6.
- 9 Murray AB, Ferguson AC, Morrison BJ. The frequency and severity of cat allergy vs dog allergy in atopic children. *J Allergy Clin Immunol* 1983;72:145-9.
- 10 Popp W, Rauscher H, Sertl K, Wanke T, Zwick H. Risk factors for sensitization to furred pets. *Allergy* 1990;45:75-9.
- 11 Arshad SH. Pets and atopic disorders in infancy. *Br J Clin Pract* 1991;45:88-9.
- 12 Dekker C, Dales R, Bartlett S, Brunekreef B, Zwanenburg H. Childhood asthma and the indoor environment. *Chest* 1991;100:922-6.
- 13 Hosein HR, Corey P, Robertson JM. The effect of domestic factors on respiratory symptoms and FEV1. *Int J Epidemiol* 1989;18:390-6.
- 14 Clifford RD, Radford M, Howell JB, Holgate ST. Prevalence of respiratory symptoms among 7 and 11 year old schoolchildren and association with asthma. *Arch Dis Child* 1989;64:1118-25.
- 15 Hong CY, Ng TP, Wong ML, Koh KTC, Goh LG, Ling SL. Lifestyle and behavioural risk factors associated with asthma morbidity in adults. *Q J Med* 1994;87:639-45.
- 16 Kuehr J, Frischer T, Karmaus W, Meinert R, Barth R, Herrman-Kunz E, et al. Early childhood risk factors for sensitization at school age. *J Allergy Clin Immunol* 1992;90:358-63.
- 17 Shatz G, Sullivan TJ, Kulczycki Jr A, Zull D, Yecies LD, Wedner HJ. Feather pillows and allergic patients [letter]. *Ann Allergy* 1982;48:59.

(Accepted 27 July 1995)

Evidence based purchasing: understanding results of clinical trials and systematic reviews

T Fahey, S Griffiths, T J Peters

Abstract

Objective—To assess whether the way in which the results of a randomised controlled trial and a systematic review are presented influences health policy decisions.

Design—A postal questionnaire to all members of a health authority within one regional health authority.

Setting—Anglia and Oxford regional health authorities.

Subjects—182 executive and non-executive members of 13 health authorities, family health services authorities, or health commissions.

Main outcome measures—The average score from all health authority members in terms of their willingness to fund a mammography programme or cardiac rehabilitation programme according to four different ways of presenting the same results of research evidence—namely, as a relative risk reduction, absolute risk reduction, proportion of event free patients, or as the number of patients needed to be treated to prevent an adverse event.

Results—The willingness to fund either programme was significantly influenced by the way in which data were presented. Results of both programmes when expressed as relative risk reductions produced significantly higher scores

when compared with other methods ($P < 0.05$). The difference was more extreme for mammography, for which the outcome condition is rarer.

Conclusions—The method of reporting trial results has a considerable influence on the health policy decisions made by health authority members.

Introduction

The randomised controlled trial is regarded as the gold standard in the assessment of healthcare interventions.¹ Its explanatory power permits qualitative conclusions about whether a treatment works and a quantitative assessment of the extent to which it works.² Purchasing organisations are now being exhorted to make health policy and purchasing decisions in terms of evidence of clinical effectiveness.^{3,4} A large proportion of such evidence comes in the form of randomised controlled trials. Research has shown that the way in which the results from controlled trials are expressed has a significant influence on physicians' willingness to prescribe drugs⁵⁻¹⁰ and patients' perception of benefit from treatment.¹¹ Whether or not the framing of the results of controlled trials has an influence on purchasing and health policy decisions has not been examined.

There are at least four ways of reporting outcomes of

Department of Public Health Medicine and Health Policy, Oxfordshire Health, Oxford OX3 9DZ
T Fahey, senior registrar in public health medicine
S Griffiths, director of public health medicine

Department of Social Medicine, University of Bristol, Bristol BS8 2PR
T J Peters, senior lecturer in medical statistics

Correspondence to:
Dr T Fahey, Department of Social Medicine, University of Bristol, Bristol BS8 2PR.

BMJ 1995;311:1056-60

clinical trials: as a relative risk reduction, absolute risk reduction, proportion of event free patients, and the number of patients who need to be treated to prevent an event.¹² Unfortunately, all too often when the results of randomised controlled trials are reported only one summary measurement of efficacy is used, most commonly the relative risk reduction. The problem with this approach is that relative risk measurements give the reader no idea of the background event rate—that is, the susceptibility of the population to the outcome of interest. Thus, when relative benefits are considerable and the outcome of interest rare, the relative risk reduction may remain high, although the absolute value of treatment will fall substantially. Quite often this problem is confounded in secondary reports and subsequent editorials that also emphasise relative differences at the expense of absolute benefits. As Feinstein states “clinicians are much impressed by the bigger numbers of relative changes than by the smaller magnitudes of the absolute changes for the same results.”¹³

With the advent of evidence based medicine, understanding the quantitative methods of data presentation from clinical trials is of major importance in terms of informed decision making in health policy.² Our aim was to assess whether health policy decisions are influenced by the way in which the results of clinical trials and systematic reviews are presented.

Methods

We sent a postal questionnaire to the 182 executive and non-executive members of health authorities, health commissions, and family health services authorities within the Anglia and Oxford regional health authorities. These authorities are responsible for the setting of health policy for their local populations and are made up of people with a broad range of skills. Executive members have responsibility for purchasing, financial, and personnel management, and directors of public health have specific training in epidemiology and include in their responsibilities the assessment of the effectiveness of medical interventions and the needs of their local population. Non-executive members, most of whom are non-medical, are appointed by the chairperson of the regional health authority on behalf of the Secretary of State to act as public representatives for their local population.

The questionnaire was constructed to present data from a single clinical trial on the efficacy of breast cancer screening¹⁴ and of a single systematic review on the efficacy of cardiac rehabilitation.¹⁵ For each programme, data were presented in four different ways in the following order: as a relative risk reduction, an absolute risk reduction, as the proportion of event free patients, and as the number of patients who needed to be treated to prevent one death.

The absolute risk reduction (absolute difference) is $(x-y)$ where x is the proportion of patients suffering the outcome of interest in the control group and y is the proportion of patients suffering the same outcome in the treatment group.¹⁶ The outcome in the control group, x , thus defines the background risk. The number of patients who need to be treated to prevent one death is the reciprocal of the absolute risk difference, $1/(x-y)$, and the relative risk reduction (relative difference) is $[(x-y)/x] \times 100$. The proportion of event free patients is the percentage of patients in each group, x and y , expressed in terms of survival rather than death. Table I shows a worked example from a systematic review that examined the efficacy of coronary artery bypass grafting¹⁷ in terms of survival benefit to illustrate the various methods of data presentation.

Each set of data was presented on the questionnaire

as the result of a different “trial” (appendix). Respondents were asked to mark an “X” along a linear scale marked from 0 (“I would not support the purchasing of this service”) to 10 (“I would strongly support the purchasing of this service”) in terms of a health authority decision to fund a breast screening programme and a cardiac rehabilitation programme. Respondents were not asked to compare the two programmes with each other. It was stated clearly that the decision to fund each programme should be taken only on the strength of the results presented from each “trial” and that costs of each alternative were the same. Respondents were asked to disregard any national or local policies with regard to the funding of either programme.

Mean scores (95% confidence interval) were calculated for each alternative presentation of results. For each programme, repeated measures analysis of variance was performed with the MANOVA procedure of spss for Windows.¹⁸ The global test for differences between the four options was carried out followed by pairwise multiple comparisons with the Bonferroni correction. All these analyses make full allowance for within person effects.

Results

Overall, 140 questionnaires were returned, three people stating that they could not give an opinion because they required further information. The response rate was therefore 75%, 79% for executive and 76% for non-executive members, and ranged from 68% to 94% across different health authorities.

Table II gives the descriptive statistics for the six pairwise comparisons between the options. The decision to fund either programme was significantly influenced by the way in which the data were presented (table III). The overall F test for differences between the four methods of data presentation was highly significant ($P < 0.001$). Reporting as a relative risk reduction produced the highest mean score to support investment in both programmes, and this was significantly higher when compared with all other methods of data presentation. In addition, the number needed

TABLE I—Method used to present the results of a randomised controlled trial or systematic review. (Calculations adapted from Guyatt et al¹⁶; example taken from Yusuf et al¹⁷)

Treatment	Outcome at 10 years		Total No of patients randomised in each group
	Death	Survival	
Coronary artery bypass grafting	350	974	1325
Medical therapy	404	921	1324
Risk of death at 10 years without therapy—medical treatment (background risk): $x=404/1324=0.305$ or 30.5%			
Risk of death at 10 years with therapy—coronary artery bypass grafting (risk with therapy): $y=350/1325=0.264$ or 26.4%			
Relative risk: $y/x=0.87$			
Relative risk reduction: $[(x-y)/x] \times 100$ or $[1-(y/x)] \times 100=13.3\%$			
Absolute risk reduction (or risk difference): $x-y=0.04$			
Event free patients is absolute percentage in each group living at 10 years: for background risk group is $(1-x) \times 100$. For therapy group is $(1-y) \times 100$.			
Individuals allocated to medical treatment=69.5%			
Individuals allocated to coronary artery bypass grafting=73.6%			
Number needed to be treated to prevent one death is $1/(x-y)=1/0.04=25$			

TABLE II—Descriptive statistics for pairwise differences between four options for each programme

Comparison of four options	Mean (SD) difference for mammography	Mean (SD) difference for cardiac rehabilitation
ARR-RRR	-4.10 (3.23)	-1.96 (2.78)
EFP-RRR	-4.12 (3.04)	-4.12 (3.04)
NNT-RRR	-2.84 (2.80)	-1.34 (2.61)
EFP-ARR	-0.022 (1.57)	-0.33 (1.74)
NNT-ARR	1.26 (2.92)	0.062 (2.40)
NNT-EFP	0.95 (2.45)	1.28 (2.83)

RRR=relative risk reduction; ARR=absolute risk reduction; EFP=event free patients; NNT=number needed to treat.

TABLE III—Mean score (%) (95% confidence interval) on two health policy issues made by health authority members when same results presented by four different methods

Method of data presentation	Mammography	Cardiac rehabilitation
Option A—relative risk reduction (RRR)	79 (76 to 83); (RRR=34%)	76 (72 to 80); (RRR=20%)
Option B—absolute risk reduction (ARR)	38 (35 to 42)†‡; (ARR=0.06%)	56 (53 to 60)†‡; (ARR=3%)
Option C—percentages of event free patients (EFP)	38 (34 to 42)†‡; (EFP=99.82% v 99.8%)	53 (49 to 57)†‡; (EFP=84% v 87%)
Option D—number needed to treat (NNT)	51 (47 to 55)†‡; (NNT=1592)	62 (58 to 66)†‡; (NNT=31)

*Higher score=stronger support for the purchasing of the programme.

Bonferroni difference means: †significant difference at $P \leq 0.05$ from relative risk reduction; ‡significant difference at $P \leq 0.05$ from number needed to treat.

to treat was significantly higher than the absolute risk reduction and the percentage event free for mammography and higher than just the latter for cardiac rehabilitation (table III). Only three respondents (all non-executive members claiming no training in epidemiology) stated that they realised that all four sets of data summarised the same results.

Discussion

This study shows that interpretation of clinical trials and systematic reviews is influenced by the method used to summarise the results. Health policy makers, like clinicians,⁵⁻⁹ are more impressed by measurements that report relative benefits compared with those entailing an index of absolute benefit. The mean estimate obtained with the relative risk reduction for both scenarios was distinctly higher than the mean obtained with all other methods. Furthermore, such differences are most extreme when the background susceptibility—in this example, death from breast cancer—is lowest (table III).

The two scenarios were chosen deliberately to reflect the fact that absolute benefits of interventions can differ considerably, even though the relative benefits may be roughly similar. The background risk of death is 85 times higher in patients who have already suffered a myocardial infarct (background risk of death 16% over three years) compared with middle aged women who are eligible to enter a breast screening programme (background risk of death from breast cancer 0.19% over seven years). Such large absolute differences in benefit are not reflected in the results of trials when relative differences are reported alone.

Consequently, it is essential that quantitative data

presented in trials and systematic reviews reflect absolute as well as relative benefits.¹⁹ This is particularly important in health policy, where background risks vary considerably according to whether interventions are aimed at populations with or without established disease. The number needed to treat possesses useful properties as a versatile way of presenting data. Calculation is relatively straightforward, background risk is incorporated into the calculation, and methods are now available to adjust it according to background risk in the target population.²⁰ Estimation of numbers needed to treat is also useful when the efficacy of different forms of interventions, such as therapeutic trials, surgical procedures, and screening programmes are compared.¹² Associated efforts, costs, and side effects of different forms of interventions can then be measured in a common currency. Furthermore, in the absence of information about the costs of alternative programmes the number needed to treat is more realistic as a proxy for cost effectiveness than is relative risk reduction.

Other commentators have suggested that the results of randomised trials and meta-analyses may require different forms of presentation when benefits at a health policy level compared with an individual level are assessed. For example, a systematic review of survival benefits from coronary artery bypass grafting on patients with stable angina expressed results in two ways.¹⁷ Firstly, as an aggregate mean survival benefit of 4.26 months at 10 years for all patients randomised to bypass grafting; and, secondly, as a disaggregated benefit (results broken down by subgroups within those randomised to bypass grafting) in which 16% of patients gained between 1.5 and 3.5 years whereas the remaining 84% had no change in life expectancy.²¹ Again, there is evidence to suggest that doctors are more impressed when the associated gain in life expectancy is expressed in a disaggregated form rather than an albeit equivalent, average figure.²² What should be recognised from all these examples^{5-10, 22} is that informed clinical and health policy decisions should be made on the basis of the full understanding and limitations of the data presented. Such decisions depend on scientific description and substantive perception, not on statistical probability.¹³

Purchasing authorities are being exhorted to increase investment in clinically effective interventions and reduce investment in interventions that evidence has shown to be ineffective.²³ With the establishment of the NHS research and development programme an increasing quantity of research is likely to be presented to policy makers in the form of clinical trials and meta-analyses.⁴ In particular, Cochrane Collaboration systematic reviews and publications from the NHS Centre for Reviews and Dissemination require an understanding of the quantitative and qualitative methods used in summarising evidence. Understanding the advantages and limitations of data presentation is fundamental to the proper interpretation of such research.

This study has shown that understanding on the part of policy makers about how results of controlled trials and systematic reviews are presented needs to be enhanced. The critical appraisal skills for purchasers (CASP) project aims to help purchasers and other decision makers develop their skills in critically appraising evidence about effectiveness.²⁴ Results from this study suggest that health authorities need to consider ways in which the critical appraisal skills of their purchasers might be improved.

We would like to thank David Sackett, professor of clinical epidemiology, Centre for Evidence Based Medicine, John Radcliffe Hospital, University of Oxford, and Dr Ruairidh Milne, consultant in public health medicine, Anglia and

Key messages

- Results of clinical trials and systematic reviews can be presented in different but equivalent ways
- Previous studies have shown that doctors' and patients' preferences for treatment are strongly influenced by the method used to present results of trials and reviews
- In this study the same results from a clinical trial and a systematic review were presented by using four different methods to health authority purchasers
- Health authority members' willingness to purchase services was influenced by the methods used to present results
- A full understanding of the ways in which the results of trials and reviews are presented is essential for informed purchasing and health policy decision making

Oxford Regional Health Authority, for their helpful suggestions on earlier drafts.

Source of funding: None.

Conflict of interest: None.

Appendix

MAMMOGRAPHY QUESTIONNAIRE

There is a proposal to offer a breast screening programme to women aged 50-64 in your health authority. There are doubts about the effectiveness of such a screening programme. We will give you four statements derived from four different randomised controlled trials published in medical journals. On the basis of each statement you should indicate how likely you are to agree to the implementation of a breast screening programme. Assume that the costs of each programme are the same. Each result was deemed to be significant.

During a seven year follow up:

- Programme A reduced the rate of deaths from breast cancer by 34%
- Programme B produced an absolute reduction in deaths from breast cancer of 0.06%
- Programme C increased the rate of patients surviving breast cancer from 99.82% to 99.88%
- Programme D meant that 1592 women needed to be screened to prevent one death from breast cancer.

CARDIAC REHABILITATION QUESTIONNAIRE

There is a proposal to offer a cardiac rehabilitation programme to people who have suffered a heart attack (myocardial infarct) in your health authority. There are doubts about the effectiveness of such a rehabilitation programme. We will give you four statements derived from four different randomised controlled trials published in medical journals. On the basis of each statement you should indicate how likely you are to agree to the implementation of a cardiac rehabilitation programme. Assume that the costs of each programme are the same. Each result was deemed to be significant.

During a three year follow up:

- Programme A reduced the rate of deaths by 20%
- Programme B produced an absolute reduction in deaths of 3%
- Programme C increased the rate of patient survival from 84% to 87%

- Programme D meant that 31 people needed to enter a rehabilitation programme to prevent one death.

- 1 Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. 2nd ed. Boston: Little, Brown, 1991.
- 2 Sackett DL, Cook RJ. Understanding clinical trials. *BMJ* 1994;309:755-6.
- 3 NHS Management Executive. *Improving clinical effectiveness*. Leeds: NHS Management Executive, 1993. (EL(93)115.)
- 4 Ham C, Hunter DJ, Robinson R. Evidence based policymaking. *BMJ* 1995;310:71-2.
- 5 Laupacis A, Sackett DL, Roberts RS. Therapeutic priorities of Canadian internists. *Can Med Assoc J* 1990;142:329-33.
- 6 Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perception of therapeutic effectiveness? *Ann Intern Med* 1992;117:916-21.
- 7 Forrow L, Taylor W, Arnold R. Absolutely relative: how research results are summarised can affect treatment decisions. *Am J Med* 1992;92:121-4.
- 8 Farkouh ME, Sauve JS, Kassam H, Lee HN, Sackett DL. Varying formats in which trial results are reported can affect clinical decision making [abstract]. *Clin Res* 1993;41:517.
- 9 Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians' willingness to prescribe. *Lancet* 1994;343:1209-11.
- 10 Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. *BMJ* 1994;309:761-4.
- 11 Malenka DJ, Baron JA, Johansen S, Wahrenberger JW, Ross JM. The framing effect of relative and absolute risk. *J Gen Intern Med* 1993;8:543-8.
- 12 Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;318:1728-33.
- 13 Feinstein AR. Invidious comparisons and unmet clinical challenges. *Am J Med* 1992;92:117-20.
- 14 Tabar L, Fagerberg G, Gad A, Baldetorp L, Holmeberg LH, Grontoft O, et al. Reductions in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;i:829-32.
- 15 Oldridge NB, Guyatt GH, Fischer ME, Rimm AA. Cardiac rehabilitation after myocardial infarction: combined experience of randomized controlled trials. *JAMA* 1988;260:945-50.
- 16 Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Working Group. Users' guides to the medical literature, 11 How to use an article about therapy or prevention, B: what are the results and will they help me in caring for my patients? *JAMA* 1994;271:59-63.
- 17 Yusuf S, Zucker D, Peduzzi P, Fisher LD, Takaro T, Kennedy WJ, et al. Effect of coronary artery bypass surgery on survival: overview of 10-year results from randomised trials by the coronary artery bypass graft surgery trialists collaboration. *Lancet* 1994;344:563-70.
- 18 SPSS. *SPSS for windows. Base system user's guide, release 6.0*. Chicago: SPSS, 1993.
- 19 Laupacis A, Naylor CD, Sackett DL. How should the results of clinical trials be presented to clinicians? [abstract]. *American College of Physicians Journal Club* 1992;May/Jun:12-4.
- 20 Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment. *BMJ* 1995;310:452-4.
- 21 Hux J, Naylor CD. Effect of coronary artery bypass surgery on survival [letters]. *Lancet* 1994;344:1223-4.
- 22 Hux JE, Levington CM, Naylor CD. Prescribing propensity: influence of life-expectancy gains and drug costs. *J Gen Intern Med* 1994;9:195-201.
- 23 Hayward J. Purchasing clinically effective care. *BMJ* 1994;309:823-4.
- 24 Institute of Health Sciences. *Critical appraisal skills for purchasers*. Oxford: IHS, 1994. (Report to the NHS Executive.)

(Accepted 25 July 1995)

Commentary: The message is in the medium

Ruairidh Milne, David Sackett

A growing number of studies have shown that the way in which evidence about a treatment's effectiveness is presented to clinicians affects how they react to it and may (arguably) lead to more rational clinical judgments. In this paper Fahey *et al* show that members of health authorities, the vast majority of whom are not medically trained, react like doctors to different ways of presenting the same information.

The authors did this by sending a questionnaire to health authority members asking how likely they would be to support the purchasing of two health care programmes, taking into account only the evidence from four "trials." In each case there was in fact only one study, but the results of that study were presented in four different ways. They found most willingness to support programmes when impact on mortality was expressed as a relative risk reduction and least support when the results were expressed with measures that combined the relative risk reduction with information about untreated patients' susceptibility to death: the absolute risk reduction and the number needed to treat.

This was clearly an artificial exercise: no alternative purchasing options were specified, no costs were included, and respondents were told to disregard national or local policies; and of course purchasing decisions are not made on the basis of evidence alone. Nevertheless, the strength of this approach was that it allowed the authors quickly and simply to assess how health authority members might react to evidence.

As well as its important central findings this paper throws up some intriguing side questions. For instance, in the case of mammography (though not cardiac rehabilitation) the numbers needed to treat produced much greater enthusiasm for purchase than did the absolute risk reduction (even though they are simply reciprocals of each other). Might this be because health authority members knew that mammography in the United Kingdom is offered mainly in the context of population screening, and so a number needed to treat of 1592 did not seem too large? It is intriguing, too, that the only people to say that they realised that all four "results" came from the same study were non-medical members.

Department of Public Health and Primary Care, Radcliffe Infirmary, Oxford OX2 6HE
Ruairidh Milne, *honorary senior clinical lecturer in public health medicine*

Centre for Evidence Based Medicine, John Radcliffe Hospital, Oxford OX3 9DU
David Sackett, *professor of clinical epidemiology*