sured, discounting of future costs and benefits, or how costs are related to outcomes.⁸

The simplest studies are those calculating cost per life saved. Two such studies have used modelling techniques applied retrospectively to trial results and produced different answers, at least partly because different perspectives on costs were used. One estimated the direct cost per life saved for a moderate infarct as \$171 000 (1987 prices) for intravenous streptokinase and \$158 000 for intravenous tissue plasminogen activator.⁹ A similar study, but limited to Medicare costs, calculated a direct cost per life saved of \$52 700 for streptokinase and \$56 900 for tissue plasminogen activator.¹⁰ Survival data can also be used for calculating cost per added life year. This has recently been reported for intravenous tissue plasminogen activator as a discounted cost per added life year of \$29 000, which was based on the results at 30 days of follow up in the trial of global utilisation of streptokinase and tissue plasminogen activator for occluded coronary arteries (GUSTO).¹¹

None of these reported studies, however, incorporate quality of life measures; such studies remain rare and vary in the instruments used to measure quality of life. Retrospective studies of intracoronary streptokinase have estimated a direct cost of \$3300 (1993 prices) for each additional quality adjusted life year based on survival at one year¹²; a subsequent study using slightly different methods estimated a discounted value of \$4000 (1993 prices) based on the results at three years of follow up in the same randomised clinical trial.¹³ Finally, studies that measure costs and quality of life prospectively in parallel with clinical trials are extremely rare. One such study of tissue plasminogen activator estimated a cost per quality adjusted life year saved of \$1000 (1993 prices) including indirect costs.⁸

The different methods used make existing research findings difficult to interpret and apply. Standardisation of

instruments to measure both clinical and economic factors is obviously a critically important next step in the evolution of useful research based information. The fact that this is difficult should not be viewed as an adequate reason for failing to take this step. On the contrary, it indicates that more research effort is needed so that interventions after myocardial infarction can be systematically evaluated, something which the NHS Standing Group on Health Technology has identified as a top priority for the NHS.⁵

ALA SZCZEPURA
Deputy director

Centre for Health Services Studies,
University of Warwick,
Coventry CV4 7AL

- 1 NHS Management Executive. *Improving the effectiveness of the NHS*. Leeds: NHS Executive, 1994.(EL(94)74.)
- 2 Hayward J. Purchasing clinically effective care. *BMJ* 1994;309:823-4.
- 3 NHS Management Executive. *Improving clinical effectiveness*. Leeds: NHS Management Executive, 1993.(EL(93)115.)
- 4 NHS Executive. *Priorities and planning guidance for 1995/96*. Leeds: NHS Executive 1994.(EL(94)55.)
- 5 Department of Health, Standing Group on Health Technology. *1994 Report*. London: Department of Health, 1994.
- 6 McMurray J, Rankin A. Recent advances in cardiology. I. Treatment of myocardial infarction, unstable angina, and angina pectoris. *BMJ* 1994;309:1343-50.
- 7 McNeil BJ. Use of claims data to monitor patients over time: acute myocardial infarction as a case study. In: Warren KS, Mosteller F. *Doing more good than harm: the evaluation of health care interventions*. *Ann NY Acad Sci* 1993;703:63-73.
- 8 Levin L-A. Thrombolytics in acute myocardial infarction. In: Szczepura AK, Kankaanpää J, eds. *Assessment of health care technologies: case studies, key concepts and strategic issues*. Chichester: Wiley (in press).
- 9 Laffel GL, Fineberg HV, Braunwald E. A cost effectiveness model for coronary thrombolysis/reperfusion therapy. *J Am Coll Cardiol* 1987;10:79-90B.
- 10 Steinberg EP, Topol EJ, Sakin JW, Kahane SN, Appel LJ, Powe NR, et al. Cost and procedure implications of thrombolytic therapy for acute myocardial infarction. *J Am Coll Cardiol* 1988;12:58-68A.
- 11 Mark DB, Naylor D, Nelson CL, Joilis JG, Clapp-Channing N, Hlatky MA. Cost effectiveness of tissue plasminogen activator relative to streptokinase in acute myocardial infarction: results from the GUSTO trial. American Heart Association conference, Atlanta, 8-11 November 1993.
- 12 Vermeer F, Simoons ML, De Zwaan C, Van ES, Verheugt FWA, Van Der Laarse A, et al. Cost benefit analysis of early thrombolytic treatment with intracoronary streptokinase. Twelve month follow up report of the randomised multicentre trial conducted by the Interniversity Cardiology Institute of The Netherlands. *Br Heart J* 1988;59:527-34.
- 13 Simoons ML, Vos J, Martens LL. Cost-utility analysis of thrombolytic therapy. *Eur Heart J* 1991;12:694-9.

Need for rigorous assessment of palliative care

Although difficult, randomised controlled trials are mandatory

See p1340

By a strange coincidence, political and professional agendas simultaneously require evidence of effectiveness. This makes good sense with finite resources. Stop ineffective interventions, and the money saved can be spent on those that are effective. The problem is that, although effectiveness is a useful word to describe what is wanted, it is difficult to define. Its meaning varies according to context.

In medicine we can at least use the gold standard of the randomised controlled trial to show that a new medical intervention is better than no intervention or than an existing intervention.¹ Long established interventions are often time honoured rather than proved by randomised controlled trials, although lack of evidence of effectiveness is not the same as evidence of lack of effectiveness. Will medical decisions not supported by randomised controlled trials continue to be purchased, and, if so, for how long? Are we wrong to continue to support these decisions? Who will investigate their effectiveness? And should one randomised controlled trial showing lack of effect outweigh the testimony of experience?² If this is hard going for single interventions how much harder is it for a service in which multiple interventions are provided?

For such services we assume that randomised controlled

trials are again the gold standard to establish effectiveness, even if the interventions that make up the service are not proved by such trials. The provision of "hospital at home" services is a good example. There are no randomised controlled trials of hospital at home. Despite this the four counties of the former Oxford Regional Health Authority will spend £1.6m on hospital at home in 1994-5. The design and conduct of a randomised controlled trial funded jointly by Northamptonshire Health Authority and Anglia and Oxford Regional Health Authority initially presented formidable problems (S Shepherd, personal communication).

This week's journal reports the failure of an attempt to evaluate a palliative care service with a randomised controlled trial (p 1340).⁵ The trial was done to "strengthen the case for continued funding." It failed because of difficulties in recruitment and problems inherent in the design. The wider issues raised by the authors are important. First is the idea that doing a randomised controlled trial in palliative care is so difficult that palliative care, and by analogy any other specialty in which it is difficult, should be "excused" from doing randomised controlled trials. The second is that the rarity of successful ran-

domised controlled trials in health services research—they quote seven adequate trials out of 1000—suggests that randomised controlled trials are not the right way to show that a service, a package of interventions, is an improvement in terms of either where it is provided or what is provided.

Should palliative care be excused from randomised controlled trials? The design of McWhinney and colleagues' trial did not make life easy. The authors had been providing their palliative care at home for 18 months before the trial began and, not surprisingly, believed that it was better than no service. Patients were randomised either to immediate support or to support delayed for one month. For those randomised to the delayed group care was meant to be provided by the family doctor, who had referred the patient in the first place. The authors measured pain, nausea, patients' quality of life, and care givers' health at the beginning of the trial and at one month. Patients expected to live more than two months were approached, yet 36 of the 146 who were randomised died within the month of the study. Others, who were not entered because they were not expected to live for two months, did so; such problems will be familiar to those who do randomised controlled trials in palliative care.

Clearly the failure of this study alone would not justify excusing palliative care from randomised controlled trials. The authors go on to raise the problems of the ethics of randomisation in palliative care in general. As they say, using patients on the waiting list as controls may be ethical in chronic stable conditions, but in this study, and in palliative care in general, randomisation to treatment versus delayed treatment is hard to justify. This does not mean that all randomisations in palliative care are unethical. Comparison of two interventions, neither of which is known to be superior, would be the classic equipoise requirement for ethical randomisation. Two constructive

options are that multicentre randomised controlled trials of particular interventions can circumvent the problem of small numbers in a single centre and that some interventions lend themselves to study by randomised controlled trials in single patients (n of 1 trials).⁴⁻⁶ In this particular case randomisation to either the full intervention or to some component might have worked better.

The provision of a service is much more subjective than the provision of an intervention. It depends more on local conditions and, indeed, on local enthusiasms. What works in one set of circumstances may not be right for another. Health care consumes a large segment of our national wealth and touches the lives of everyone. The time has come when we should put an end to the casual use of "effectiveness" as an easy way out and begin to take a harder line. If you say that it is effective then prove it. As Archie Cochrane wrote more than 20 years ago, "There will be a marked reduction in the use of ineffective remedies and of effective remedies used inefficiently."⁷

HENRY MCQUAY
Clinical reader in pain relief
ANDREW MOORE
Consultant biochemist

Pain Relief Unit,
The Churchill,
Oxford OX3 7LJ

- 1 Advisory Group on Health Technology Assessment. *Assessing the effects of health technologies*. London: Department of Health, 1992.
- 2 Eddy DM. Three battles to watch in the 1990s. *JAMA* 1993;270:520-6.
- 3 McWhinney IR, Bass MJ, Donner A. Evaluation of a palliative care service: problems and pitfalls. *BMJ* 1994;309:1340-2.
- 4 Guyatt G, Sackett D, Taylor WD, Chong J, Roberts R, Pugsley S. Determining optimal therapy—randomised trials in individual patients. *N Engl J Med* 1986;314:889-92.
- 5 McQuay HJ. N of 1 trials. In: Max MB, Portenoy RK, Laska EM, eds. *The design of analgesic clinical trials*. New York: Raven Press, 1991:179-92. (Advances in pain research and therapy Vol 18.)
- 6 March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P. n of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. *BMJ* 1994;309:1041-6.
- 7 Cochrane AL. *Effectiveness and efficiency. Random reflections on health services*. London: Nuffield Provincial Hospitals Trust, 1972:84.

Changes in semen and the testis

All the trends suggest a decline in the health of the male reproductive system

Infertility currently affects at least one man in 20, but we are unable to assess whether it is becoming more prevalent because we have extremely poor past records. Male reproduction has been given little attention scientifically and in medical practice. However, our recent meta-analysis in the *BMJ* of 61 papers on semen quality, which clearly pointed to a decrease over the past 50 years,¹ has been repeatedly quoted in the *BMJ*,²⁻⁴ *Science*,⁵ and the popular media.

We have already responded to the various technical criticisms, several of which concerned the known limitations of meta-analysis.⁶⁻⁸ Though the data for 1970-90 were compatible with a decrease as well as with no change or an increase in semen quality, the cautious general conclusion is that a real decline in semen quality did occur over the full period 1940-90. Three recent reports have found that semen quality has declined among candidate semen donors during the past 20 years.⁹⁻¹¹ Two of these studies not only included data on sperm count but also showed that important measures of quality—morphology of sperm heads and motility—had also deteriorated.^{9,10}

These changes are all the more worrying since the quality of "normal" human semen is already very poor when compared with that of other mammalian species—

more than half of our sperm usually have one or more morphological abnormalities compared with less than 5% in some other species—and several studies have shown a relation between abnormal sperm heads and infertility.¹² Whether the prevalence of abnormalities in human sperm has always been so high is not known; some recent reports indicate that a change might have occurred.^{9,10} The relatively high number of abnormal sperm may explain in part our low fecundity compared with that of most mammals.

The evidence of a decline in the quality of semen is not the only indicator that the human testis is at risk. For a generation or more, cancer of the testis has been becoming more common.¹³ This is a disorder of young men and is associated with a high rate of other abnormalities of the testis, including undescended testis and poor semen quality. One quarter of men with unilateral cancer of the testis have severe impairment of the contralateral testis, including the "Sertoli cell only" pattern and spermatogenic arrest. Furthermore, the incidence of both hypospadias and undescended testes has been reported to be rising in the general population. The evidence of declining semen quality should be seen in the light of these trends in other reproductive disorders in men.¹²