Letters

Website: **bmj.com**
Email: **letters@bmj.com**

Routine audit is an ethical requirement of screening

Editor—Cervical screening saves about 1300 lives each year in England and Wales.[1] We regard audit as an essential part of the screening programme and urge health authorities to continue this activity despite recent concerns about using patient information without informed consent.

Poor quality screening is ineffective and may do more harm than good. Women screened in the NHS can expect a high quality service in which smears are properly taken and read to a high standard and the results stored to ensure appropriate management. Audit is part of the quality assurance that is integral to the screening that each woman receives.

Since 1988 records of every smear test have been entered on to health authority databases. The dates and results of all smear tests are linked to NHS numbers since they are used to determine the timing of future tests. The clinical value of these databases is enormous. Before they existed, coverage was poor and follow up of women with abnormal results was often inadequate.

To evaluate the effectiveness of a screening programme and to identify its strengths and weaknesses screening histories sampled from the entire target population must be audited. This enables rational decisions to be made about modifications on issues such as quality, screening interval, target age groups, the need for an improved screening test, the importance of improving failsafe mechanisms, and the potential gain from improved coverage. Reliable audits cannot depend on consenting women but must be representative of the whole population. Analyses based only on consenting women are likely to be biased and misleading.

Such an audit has been running under the auspices of the national screening programmes since 1992. After linking screening histories to women diagnosed with cervical cancer, health authorities have sent anonymised records from over 7500 women (including 2500 with cancer) to the Imperial Cancer Research Fund for analysis. The data are stored on a secure computer system in password protected directories and are made public only in aggregated form. No one at the fund knows the identity of the women whose screening histories are held in this audit.

We believe that routine audit is an ethical requirement of a screening programme. The benefit in terms of cancer prevention is sufficiently great to warrant the secretary of state making regulations in accordance with clause 68 of the Health and Social Care Bill, and we urge him to do so.

The issues go far beyond cervical screening. Disease prevention and health promotion activities must be audited for the future public health of the country. Only in this way can we ensure that these initiatives are achieving their goals and giving the best protection possible.

Peter Sasieni *coordinator, cervical screening audit*
Jack Cuzick *head of mathematics, statistics, and epidemiology*
Imperial Cancer Research Fund, London WC2A 3PX

On behalf of 13 other authors: Peter Boyle, chairman of prevention and control, Imperial Cancer Research Fund; Penny Craddock, chairman, WNCCC—Cancer Aware; Trevor Hince, director, scientific department, Cancer Research Campaign; Henry Kitchener, president, British Society of Colposcopy and Cervical Pathology; John Lilleyman, president, Royal College of Pathologists; James McEwen, president, Faculty of Public Health Medicine; Rebecca Miles, senior manager, National Cancer Alliance; Monica Roche, chairman, UK Association of Cancer Registries; Maurice Slevin, chairman, CancerBACUP; Martin Vessey, emeritus professor of public health, University of Oxford; Nicholas Wald, editor, *Journal of Medical Screening*; Nichola Wilkins, chief executive, Royal Institute of Public Health and Hygiene and Society of Public Health; Nicholas Young, chief executive, Macmillan Cancer Relief.

1 Sasieni PD, Adams J. Effect of screening on cervical cancer mortality in England and Wales: analysis of trends using an age-period-cohort model. *BMJ* 1999;318:1244-5.

Effect of receiving a heart transplant

Surely it is too late for a randomised controlled trial

Editor—Rigorous evaluation of surgical procedures is important. The comparative outcomes and clinical profiles in transplantation (COCPIT) study, reported by Deng et al, called for a randomised controlled trial of heart transplantation based on comparing postoperative and waiting list survival in a single year in Germany.[1]

The German registry's 12 month postoperative survival rate of 71% was considerably lower that that of patients at Papworth Hospital in Cambridge who have undergone transplantation since 1990 (83%) and that reported by the International Society for Heart and Lung Transplantation (82% for patients who have undergone transplantation since 1995).[2] This brings into question the generalisability of the results beyond Germany. Moreover, 12 month follow up is inadequate, as risks after transplantation are greatest in the first year. The risks from end stage heart failure are cumulative.

Deng et al claim that some patients listed to receive transplants are not sick enough to derive survival benefit from the procedure. The difficulty lies in identifying the group for which equipoise exists, at least in terms of survival. With increasing waiting lists and higher proportions of patients receiving transplants who are in United Network for Organ Sharing status I, the demand of the high risk group for donated organs may leave few organs for the marginal candidates, for whom randomisation may be appropriate.

Peak oxygen uptake is an important prognostic factor in the heart failure survival score, which is widely used by transplant centres.[3] However, it was available for only 16% of the patients in the study reported by Deng et al; substituting the mean for missing values is questionable. This casts doubt on the value of the score, which is unvalidated in patients taking β blockers or as a prognostic indicator in transplantation.[4]

Over the past 10 years compelling evidence has emerged that heart trans-

Advice to authors

We prefer to receive all responses electronically, sent either directly to our website or to the editorial office as email or on a disk. Processing your letter will be delayed unless it arrives in an electronic form.

We are now posting all direct submissions to our website within 24 hours of receipt and our intention is to post all other electronic submissions there as well. All responses will be eligible for publication in the paper journal.

Responses should be under 400 words and relate to articles published in the preceding month. They should include ≤ 5 references, in the Vancouver style, including one to the BMJ article to which they relate. We welcome illustrations.

Please supply each author's current appointment and full address, and a phone or fax number or email address for the corresponding author. We ask authors to declare any competing interest. Please send a stamped addressed envelope if you would like to know whether your letter has been accepted or rejected.

Letters will be edited and may be shortened.

bmj.com
letters@bmj.com

plantation confers considerable improvements in health related quality of life. For example, in 1990, using the Nottingham health profile, we showed considerable improvements in all physical and psychological dimensions,[5] with improvements maintained to five years in survivors. The evidence of improved quality of life has been steadily mounting from other centres.

The comparative outcomes and clinical profiles in transplantation study has raised the importance of careful selection of candidates for heart transplantation, which should be done in large, experienced centres. With mounting evidence of improved survival of transplant recipients and clinically important improvements in health related quality of life, we are surely too late for a randomised controlled trial.

D K Satchithananda *transplant research fellow*
S C Stoica *transplant research fellow*
J Parameshwar *consultant transplant physician*
S R Large *consultant surgeon*
J Wallwork *director of transplantation*
Transplant Unit, Papworth Hospital NHS Trust, Cambridge CB3 8RE

L D Sharples *senior statistician*
MRC Biostatistics Unit, Cambridge CB2 2SR
Linda.Sharples@papworth-tr.anglox.nhs.uk

1 Deng MC, De Meester MJ, Smits JMA, Heinecke J, Scheld HH on behalf of the Comparative Outcome and Clinical Profiles in Transplantation (COCPIT) Study Group. Effect of receiving a heart transplant: analysis of a national cohort entered on to a waiting list, stratified by heart failure severity [with commentary by T Treasure, A Murday]. *BMJ* 2000;321:540-5. (2 September.)
2 Hosenpud JD, Bennett LE, Keck BM, Fiol B, Boucek MM, Novick RJ. The registry of the International Society of Heart and Lung Transplantation: sixteenth official report–1999. *J Heart Lung Transplant* 1999;18:611-26.
3 Mancini D, LeJemtel T, Aaronson K. Peak VO_2: a simple yet enduring standard. *Circulation* 2000;101:1080-2.
4 Aaronson KD, Schwartz JS, Chen TM, Wong KL, Goin JE, Mancini DM. Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation* 1997;95:2660-7.
5 Caine N, Sharples LD, English TAH, Wallwork J. Prospective study comparing quality of life before and after heart transplantation. *Transplant Proc* 1990;22: 1437-9.

Peak exercise oxygen consumption is important predictor of outcome

EDITOR—We agree with Deng et al, who report the comparative outcomes and clinical profiles in transplantation (COCPIT) study, that only patients with a predicted high mortality should be listed for heart transplantation.[1] But it is important to look closely at patients' characteristics, particularly the oxygen consumption at peak exercise.

Peak exercise oxygen consumption is an important predictor of outcome in patients with heart failure[2 3] and an important factor in the heart failure survival score statistical model,[4] used in Deng et al's study. Mancini et al showed that patients with a peak exercise oxygen consumption of <14 ml/kg/min had a significantly higher mortality than patients with a peak exercise oxygen consumption of >14 ml/kg/min.[2] In Deng et al's study, however, the mean peak exercise oxygen consumption was 15.8 ml/kg/min. This variable was available for only 139 of the 889 patients, which leads to several questions. Firstly, this group of patients would seem to be relatively well and so would not be predicted to have a high mortality; perhaps many of them should not have been listed for organ transplantation. Secondly, why was this important predictor of outcome presented for only 16% of the patients?

The accurate assessment of patients for heart transplantation is critical for deriving mortality benefit from this procedure. Failure to use accepted tests such as peak exercise oxygen consumption may result in inappropriate listing of patients with heart failure for transplantation.

Guy A MacGowan *assistant professor of medicine*
Srinivas Murali *associate professor of medicine*
Section of Heart Failure and Transplantation Cardiology, Cardiovascular Institute of the University of Pittsburgh Medical Center, Pittsburgh, PA 15213, USA
macgowanga@msx.upmc.edu

1 Deng MC, De Meester MJ, Smits JMA, Heinecke J, Scheld HH on behalf of the Comparative Outcome and Clinical Profiles in Transplantation (COCPIT) Study Group. Effect of receiving a heart transplant: analysis of a national cohort entered on to a waiting list, stratified by heart failure severity [with commentary by T Treasure, A Murday]. *BMJ* 2000;321:540-5. (2 September.)
2 Mancini DM, Eisen H, Kussmaul W, Mull R, Edmunds LH, Wilson JR. Value of peak exercise oxygen consumption for optimal timing of cardiac transplantation in ambulatory patients with heart failure. *Circulation* 1991;83:778-86.
3 Mancini D, LeJemtel T, Aaronson K. Peak VO_2: a simple yet enduring standard. *Circulation* 2000;101:1080-2.
4 Aaronson K, Schwartz JS, Chen T, Wong KL, Goin JE, Mancini D. Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation* 1997;95:2660-7.

Research target should be stratification procedures and mechanisms of death

EDITOR—Deng et al have produced important contemporary observations in cardiac transplantation.[1] Cardiac transplantation is about avoiding death in the short term (months), yet their concluding emphasis on the need for a controlled trial of this treatment might be seen as a criticism of the overall benefits of this service.

Patients who are listed for transplant operations should have no other option. Coming to this conclusion is an inexact science, even with all available individual observations. The independent risk indicators are derived from population studies of longitudinal survival; assessment of individual risk is statistically and practically impossible. There is often a remarkable and diverse range of opinions on when to list a "typical" case of end stage disease for transplantation, with a glaring lack of evidence to guide day to day clinical practice.

The suggestion that patients with more severe variables of advanced disease benefit more from a high risk treatment than do those with less severe indicators would apply to any treatment, not just heart transplantation. To suggest that some of the decisions in patients with lower risk scores are erroneous is an inappropriately harsh comment on current clinical practice.

Two issues require more research: we need a better understanding of the complexities of death in heart failure, and more knowledge of how to defer this event through the selection of high risk treatment strategies. No one would ever suggest that transplantation is a cure, but it should always lead at least to a longer life with a different form of illness.

Stratification techniques such as those summarised in the paper fail to predict which individual patients will die quickly as well as suddenly (the two are not the same) and which will live on for years despite having a low left ventricular ejection fraction. Yet where is the call to investigate this? This should be the topic for research, not questioning the efficacy of a last line of treatment.

In the United Kingdom, and I suspect in Germany too, care and sensitivity are required when dealing with our limited transplantation infrastructures. Criticism of patterns of activity that are directed entirely at the individual patient's best interests and based on the best available individual patient data is not only unhelpful but also potentially destructive. It takes little insight to generate an inappropriate headline in a daily newspaper, leaving this vulnerable group of patients feeling even worse than before about their prospects. I am sure that this was not the intention of the authors.

Robert J MacFadyen *consultant cardiologist*
Cardiac Unit, Highland Acute Hospitals, Inverness IV2 3UJ
robert.macfadyen@raigmore.scot.nhs.uk

1 Deng MC, De Meester JMJ, Smits JMA, Heinecke J, Scheld HH on behalf of the Comparative Outcome and Clinical Profiles in Transplantation (COCPIT) Study Group. Effect of receiving a heart transplant: analysis of a national cohort entered on to a waiting list, stratified by heart failure severity [with commentary by T Treasure, A Murday]. *BMJ* 2000;321:540-5. (2 September.)

Authors' reply

EDITOR—We agree that two circumstances impose important limitations on our study: the incompleteness of data and the use of the original heart failure survival score for risk stratification derived in a different group.[1]

Data collection for the cohort occurred after the enactment in December 1996 of the German transplantation law requiring all transplant doctors to define necessity, urgency, and likelihood of success of organ transplantation by adequate data collection and analysis; the heart failure survival score was published in July 1997. The reason for using the score was that it predicted what it was supposed to predict: a high risk of dying without a transplant.

Two issues require clarification: why only 139 of 889 peak oxygen values were available, although this is the most widely accepted predictor,[2] and why the entire transplant cohort had a worse survival than either the Papworth group or the International Society for Heart and Lung Transplantation registry group[3] without an apparently different risk profile. The heart committee of the German Transplantation Society has thus initiated a national auditing process similar to that in the United Kingdom.

The question on the generalisability of these data to any other national cohort would require this study to be repeated in other countries. The Eurotransplant Inter-

national Foundation has already paved the way for starting the Eurotransplant heart COCPIT (comparative outcomes and clinical profiles in transplantation) study.

Another important criticism concerns the short duration of observation. A potential benefit in the medium and low risk group might indeed become apparent only in the second year after the intervention. This long term assessment would require mathematical modelling of the transitionary aspects of risk status.

Our study was a first attempt to generate outcome data in contemporary cardiac transplantation based on a consensus of all participating centres in a national cohort and heart failure risk stratification. Without this study, survival data after listing in Germany would not have been known; adequate auditing steps can now be taken.

Our study has four implications: a rigorous (inter)national auditing process should be set up; consensus is needed on the cardiac transplant centres required; similar studies should be repeated in other countries; and the survival benefit of cardiac transplantation should be tested more rigorously. The first two points are currently being addressed in Germany: every high urgency request for cardiac transplantation is being audited by a panel of three experts, and decisions are subjected to review by the heart committee every six months. The last two points will require a dedicated and networked research effort of national and international organisations involved in management of advanced heart failure.

Mario C Deng *director of cardiac transplantation research*
Heart Failure Center, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA
dengmario@hotmail.com

Johan M J De Meester *former head, medical affairs*
Jacqueline M A Smits *senior biostatistician*
Eurotransplant Foundation, PO Box 2304, 2301 CH Leiden, Netherlands

Joachim Heinecke *senior biostatistician*
Institute for Biomathematics, Münster University, D-48129 Münster, Germany

Hans H Scheld *professor*
Department of Cardiothoracic Surgery, Münster University

1 Aaronson KD, Schwartz JS, Chen TM, Wong KL, Goin JE, Mancini DM. Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation* 1997;95:2660-7.
2 Mancini D, LeJemtel T, Aaronson K. Peak VO_2: a simple yet enduring standard. *Circulation* 2000;101:1080-2.
3 Hosenpud JD, Bennett LE, Keck BM, Boucek MM, Novick RJ. The registry of the International Society for Heart and Lung Transplantation: seventeenth official report—2000. *J Heart Lung Transplant* 2000;19:909-31.

Debate about blood pressure and epistaxis will continue

Editor—Manfredini et al in showed a circadian variation in the onset of epistaxis in their paper.[1] Epistaxis is the second most common cause of spontaneous bleeding. Of the two categories of epistaxis, mundane and severe, the mundane, usually anterior epistaxis, is the more common. Manfredini et al did not, however, report on the subgroups of epistaxis.

Sixty per cent of patients may experience at least one episode of epistaxis during their lifetime. Eighty per cent of epistaxis occurs in Kiesselbach's plexus, a vascular network in the anterior portion of the nasal septum. Beran et al reported that common colds, stress, or tiredness were often experienced before the occurrence of the nosebleeds. The blood pressure distribution of habitual nosebleeders did not differ from that of the population samples used for comparison.[2] This is in concordance with the data of Lubianca Neto et al, who also could not establish a definite association between blood pressure and history of adult epistaxis in hypertensive patients, although they found a link to left ventricular hypertrophy.[3] The evidence for an association of duration of hypertension and left ventricular hypertrophy with epistaxis suggests that epistaxis might be a consequence of long term hypertension. They also observed the presence of enlarged vessels at rhinoscopy in hypertensive patients with a history of epistaxis. Herkner et al document significantly higher blood pressure values in epistaxis patients compared with controls.[4]

The discussion on blood pressure and epistaxis will continue. However, ear, nose, and throat surgeons see that the central task in dealing with epistaxis is to differentiate between anterior and posterior origins of bleeding, and authors of further studies have to include the localisation of the bleeding, as Padgham et al did.[5] They found a positive correlation between hypertension and bleeding from the middle meatus, but not with the severity of bleeding.

Andreas F P Temmel *senior registrar*
Christian Quint *senior registrar*
Josef Toth *senior registrar*
Department of Otorhinolaryngology, University of Vienna, A-1090 Vienna, Austria

1 Manfredini R, Portaluppi F, Salmi R, Martini A, Gallerani M. Circadian variation in onset of epistaxis: analysis of hospital admissions. *BMJ* 2000;321:111. (4 November.)
2 Beran M, Petruson B. Occurrence of epistaxis in habitual nose-bleeders and analysis of some etiological factors. *ORL J Otorhinolaryngol Relat Spec* 1986;48:297-303.
3 Lubianca Neto JF, Fuchs FD, Facco SR, Gus M, Fasolo L, Mafessoni R, et al. Is epistaxis evidence of end-organ damage in patients with hypertension? *Laryngoscope* 1999;109:1111-5.
4 Herkner H, Laggner AN, Mullner M, Formanek M, Bur A, Gamper G, et al. Hypertension in patients presenting with epistaxis. *Ann Emerg Med* 2000;35:126-30.
5 Padgham N. Epistaxis: anatomical and clinical correlates. *J Laryngol Otol* 1990;104:308-11.

Principal variable is not what it seems in league tables

Editor—Sir Brian Jarman's analysis of hospital death rates with "Dr Foster's guide to better health" (*Sunday Times*) may serve to improve the quality of hospital care—indirectly.[1] The principal dependent variable is, however, not what it seems, even after adjustment for age, sex, diagnosis, emergency admission, and length of stay, so that like is not compared with like. Rates derived from hospital episode statistics, deaths per 1000 finished consultant episodes, almost defy interpretation, because the denominators are episodes, not patients. Although this analysis selects a subset of episodes that end in discharge or death, the denominators represent admissions, not people. Fairer measures of hospital performance are based on 30 day deaths per 100 000 population.[2 3]

The first conclusion of the study should read that the number of hospital episodes (or admissions) has increased by approximately 2.6% annually. The numbers of deaths have remained nearly constant. It is only a consequence of increased activity that "episode fatality rates" seems to have fallen.

The second main observation compares episode fatality rates with the ratio of doctors to beds, a ratio of two provision measures: (hospital) doctors and (acute hospital) beds per 100 000 population. It would be preferable to examine relations with these two measures of provision independently. High ratios of doctors to beds are found in tertiary centres, and low episode fatality rates in such hospitals could be an artefact of denominator inflation: more doctors in more specialties so that one patient and one illness appears as more than one episode in more than one specialty.

The third main observation, association with provision in general practice, may be true yet have little to do with the quality of hospital care, if districts and communities of high provision have appropriate alternatives for care of the dying, at home or in hospices, in the final days, after curative treatment has been abandoned. There are many factors outside hospital that affect hospital death rates even after adjustments as in this analysis, including admission and discharge policies and care in the community.[4]

Dr Foster's guide, effectively a full league table, may help health professionals and managers to identify weaknesses, where weaknesses occur, and improve services more than they alarm patients.[5] It is to be hoped that poor ratings may not be improved by the simple expedient of denominator adjustment. It should, however, not be forgotten that, across most of the country, patients do not have choice; when ill we go to "our" local hospital.

Robert West *reader in epidemiology*
University of Wales College of Medicine, Cardiff CF4 4XN

1 Jarman B, Gault S, Alves B, Hider A, Dolan S, Cook A, et al. Explaining differences in English hospital death rates using routinely collected data. *BMJ* 1999;318:1515-20.
2 Capewell S, Kendrick S, Boyd J, Cohen G, Juszczak E, Clarke J. Measuring outcomes: one month survival after acute myocardial infarction in Scotland. *Heart* 1996;76:70-5.
3 Birkhead J, Goldacre M, Mason A, Wilkinson E, Amess M, Cleary R, eds. *Health outcomes indicators, myocardial infarction: report of writing group of Department of Health.* Oxford: Health Outcomes Development, 1999.
4 Rosen M, West RR. *Urgent and emergency admission to hospital.* London: HMSO, 1995.
5 West RR. Performance guides raising the standard—indirectly? *J Pub Health Med* 1997:19:361-3.

Reduction of postoperative mortality and morbidity

Little information was given on inclusion criteria

EDITOR—Rodgers et al report a meta-analysis of 141 trials comparing general anaesthesia with neuraxial blocks.[1] They conclude that their data should result in more widespread use of spinal or epidural anaesthesia. The challenge for clinicians is deciding which of their patients (if any) these results apply to, but Rodgers et al provided little information about the inclusion criteria for the trials examined. The applicability of a meta-analysis is difficult to assess when heterogeneous patient groups are combined. Also, although a spinal or epidural anaesthetic might be reasonably standard, there are many general anaesthetic agents and these may not be comparable.

Figure 1 of the meta-analysis shows that four trials contributed 31 deaths to the overall mortality difference of 41.[2–5] In those four studies the mortality from general anaesthesia ranged from 8% to 27%, compared with 3.1% for all trials combined.

We wonder whether any information was collated on antithrombosis prophylaxis. In 2001 most patients at risk of venous thrombosis having a general anaesthetic will receive prophylaxis including anticoagulation drugs at low doses. Of the four trials referred to above, three are over 15 years old, and at least one specifically excluded patients receiving low dose anticoagulation.[3]

Figure 2 of the meta-analysis indicates an apparent benefit of neuraxial block in orthopaedic patients.[1] Mortality after vascular, urological, and general surgery showed no significant difference. We therefore question the conclusion of Rodgers et al that their result is applicable to all surgical patients. In the vascular group (the only non-orthopaedic group that approached significance) there was a difference of eight deaths. That difference would be reduced to just one death by eliminating a single trial in which the mortality after general anaesthesia was 18%.[5] We would be interested in the views of Rodgers et al on the relevance of that trial to institutions where mortality may be much lower.

It could be a mistake to apply the conclusions of a meta-analysis to any particular patient if that patient's characteristics are different to those in the initial trials or if aspects of their management differ. From the data presented, we could not come to the same conclusions as the authors.

Timothy J McCulloch *staff specialist anaesthetist*
John A Loadsman *staff specialist anaesthetist*
Department of Anaesthetics, Royal Prince Alfred Hospital, Camperdown, New South Wales 2050, Australia

1 Rodgers A, Walker N, Schug S, McKee A, Kehlet H, van Zundert A, et al. Reduction of postoperative mortality and morbidity with epidural or spinal anaesthesia: results from overview of randomised trials. *BMJ* 2000;321:1493. (16 December.)
2 McLaren AD, Stockwell MC, Reid VT. Anaesthetic techniques for surgical correction of fractured neck of femur: a comparative study of spinal and general anaesthesia in the elderly. *Anaesthesia* 1978;33:10-4.
3 Valentin N, Lomholt B, Jensen JS, Hejgaard N, Kreiner S. Spinal or general anaesthesia for surgery of the fractured hip? A prospective study of mortality in 578 patients. *Br J Anaesthesia* 1986;58:284-91.
4 McKenzie PJ, Wishart HY, Smith G. Long-term outcome after repair of fractured neck of femur. Comparison of subarachnoid and general anaesthesia. *Br J Anaesthesia* 1984;56:581-5.
5 Borovskikh NA, Lebedev LV, Strashkov VI, Vinogradov AT. [Comparative evaluation of the effectiveness of epidural anesthesia with spontaneous respiration and general anesthesia in aorto-femoral bifurcation shunt]. *Vestnik Khirurgii Imeni i-i-Grekova* 1990;145:95-8.

Research into modern anaesthesia techniques and perioperative medicine is needed

EDITOR—Rodgers et al in their meta-analysis to resolve one of the more contentious issues in anaesthesia over recent years—namely, whether there is there any advantage of regional over general anaesthesia.[1] From the results they conclude that regional techniques (spinal or epidural anaesthesia) reduce postoperative mortality and decrease the incidence of other serious complications such as pneumonia and pulmonary embolism (although the effects of regional anaesthesia on myocardial infarction and renal failure were inconclusive).

Rodgers et al suggest that these data support the more widespread use of regional anaesthetic techniques. The data have been meticulously gathered and researched from a large number of trials but we would like to point out a few areas of concern.

- Only 56 trials (40%) looked at outcome. Thirteen of these (23%) followed patients up for a period lasting more than 30 days, while 19 (34%) followed patients up for less than seven days
- One hundred and sixteen (82%) of the 141 papers that met the inclusion criteria were published before 1990. This means that data from these studies are now at least 10-12 years old
- Anaesthetic techniques, equipment, and drugs have changed quite dramatically in recent years, and, therefore, studies predating these advances may have lost some of their relevance
- Use of some of the older volatile agents has declined, and newer agents with fewer cardiovascular side effects are now in widespread use
- Stopping cardiac drug treatment preoperatively is no longer recommended
- Heparin prophylaxis for venous thrombosis and pulmonary embolism is now much more common,[2] particularly since the introduction of low molecular weight heparins
- Ten trials looked at outcome in vascular surgical patients (a group known to be at increased risk of perioperative complications), eight of which were published after 1990, and there were more myocardial infarctions and cardiac deaths in the patients who had received a regional anaesthetic.

We broadly agree with the conclusion of Rodgers et al. Morbidity and mortality seem to be reduced in patients receiving regional anaesthesia with a trend towards reduced cardiac and renal complications. However, the paucity of recent data in which hard outcome measures are assessed indicates the need for more research directed towards understanding the impact of modern anaesthesia techniques and perioperative medicine on patient outcomes.

Helen Higham *research fellow*
helen.higham@nda.ox.ac.uk

Pitabas Mishra *specialist registrar in anaesthetics*
P Foëx *professor of anaesthetics*
Nuffield Department of Anaesthetics, John Radcliffe Hospital, Headington, Oxford OX3 9DU

1 Rodgers A, Walker N, Schug S, McKee A, Kehlet H, van Zundert A, et al. Reduction of postoperative mortality and morbidity with epidural or spinal anaesthesia: results from overview of randomised trials. *BMJ* 2000;321:1493. (16 December.)
2 Francis RM, Brenkel IJ. Survey of use of thromboprophylaxis for routine total hip replacement by British orthopaedic surgeons. *Br J Hosp Med* 1997;57:427-31.

Authors' reply

EDITOR—McCulloch and Loadsman raise issues about applicability of meta-analyses of heterogeneous patient populations, implying that generalisation should be restricted to patients closely similar to those in the included trials. Because of the large number of trials, only limited data on eligibility criteria could be published, even in the web site version. The key issue, however, aiding generalisibility is not "representativeness," but consistency between different trials, especially if observed across heterogeneous patient groups.[1] The proportional effects of neuraxial blockade were broadly consistent, justifying the pooling process. Absolute risk reductions therefore increased with increasing event rates. For example, neuraxial blockade reduced deep vein thrombosis by 44% (SE 11) in trials that employed screening and 46% (22) in other trials, and so absolute differences were greater in screening versus other trials (12% *v* 0.5%, respectively). Therefore, trials with most events contributed most of the net difference in events. However, treatment effects were not clearly restricted to such trials—for example, in trials that observed more than or less than 10 deaths, mortality reductions were 33% (13) and 27% (20) respectively.

Most patient groups have lower mortality than those in the meta-analysis. But there were clear reductions, overall and in several different surgical groups, for other important outcomes such as thromboembolism, pneumonia, and bleeding. For these outcomes, it seems appropriate to require very good direct evidence of lack of benefit before safely concluding that neuraxial blockade is not effective in some particular group.[3] We did not observe such evidence.

Applying trial results to individual patients should ideally entail combining a typical proportional reduction from a meta-analysis with a patient's estimated absolute risk.[2] An updated meta-analysis with data from individual participants could improve estimates of proportional reductions and more reliably identify any subgroup effects. A key challenge for the clinician remains estimating a patient's absolute risk, however, since this is likely to vary substantially more

than any subgroup differences in proportional effects of neuraxial blockade.

Higham et al point out correctly that most trials did not report clinical outcomes. But we aimed to collect data from all trials, irrespective of their original aims. Length of follow up inevitably varies between trials, and we separately analysed events within 30 days (usually within 20) and deaths after 30 days.

Both letters point out that thromboprophylactic practices have changed in recent decades, but we excluded trials in which they were systematically different between the randomised groups. Concomitant anaesthetic practice has also changed, but the overall reduction in mortality was 28% (13) for trials published before 1990 and 35% (21) for trials published in the 1990s.

Anthony Rodgers *codirector*
Natalie Walker *research fellow*
Clinical Trials Research Unit, Department of Medicine

Stephan Schug *professor*
Division of Anaethesiology
University of Auckland, Private Bag 92019, Auckland 1, New Zealand

Andrew McKee *consultant anaesthetist*
Department of Anaesthetics, Green Lane Hospital, Epsom, Auckland 1003

Andre van Zundert *consultant anaesthetist*
Department of Anesthesiology, Intensive Care and Pain Therapy, Catharina Hospital, NL-5623 EJ Eindhoven, Netherlands

1 Rothman KJ, Greenland S. *Modern epidemiology*. 2nd ed. Philadelphia: Lippincott-Raven, 1998.
2 Glasziou PP, Irwig LM. An evidence based approach to individualising treatment. *BMJ* 1995;311:1356-9.
3 Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity. I. Clinical trials. *Lancet* 2001;357:373-80.

Dual blockade of renin-angiotensin system

Data do not support claimed benefit of combination over single treatment

Editor—Mogensen et al interpret their results as showing that the combination of lisinopril plus candesartan was more effective than either agent alone at lowering blood pressure and reducing the urinary albumin:creatinine ratio over 24 weeks in patients with type 2 (non-insulin dependent) diabetes.[1] They state, "our results … support this new and potentially highly beneficial therapeutic approach for the prevention of diabetic renal and vascular disease."

The authors' interpretation of the renoprotective effect of the combination and its superiority in lowering blood pressure does not seem to be supported by the results. Several methodological issues in particular limit their interpretation.

Firstly, the authors state in the abstract that "the reduction in urinary albumin:creatinine ratio with combination treatment … was greater than with either candesartan … or lisinopril." Table 4, however, shows that the difference in the albumin:creatinine ratio between lisinopril and the combination was not significant ($P > 0.20$); they did not show that the combination was more effective than lisinopril alone. The P value for the difference between candesartan and the combination was 0.04. As numerous statistical tests were done and no correction for multiple testing was applied we question whether a P value of 0.04 is significant.

Secondly, the authors report that the combination produced increases in creatinine and urea concentrations and a decrease in creatinine clearance at 24 weeks. These changes at best should be going in the opposite direction, or at least be no different if the combination is more renoprotective.

Thirdly, from a methodological perspective, two factors need explaining.

Firstly, patients received combination treatment only during weeks 12-24. Inexplicably, the authors compared the effects of the combined treatment (at 24 weeks) with baseline data (week 0). Secondly, 53 patients (27%) who had been enrolled in the trial were excluded from these comparisons because their diastolic pressures had been reduced to <80 mm Hg at week 12 with single treatment.

In essence, the authors compared patients who received the combination (including carry-over effects from 12 weeks of single treatment) with patients who had the least response to single treatment. This is a biased comparison, and the authors should have compared single and combination treatment from 12 to 24 weeks, not from baseline. Presumably, the 53 patients were excluded because it was deemed unsafe or unnecessary to reduce their blood pressures further.

It would be of concern if clinicians concluded from this study that it is beneficial to use the combination at the outset of treatment.

James McCormack *associate professor*
jmccorma@interchange.ubc.ca

Marc Levine *professor*
Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

Competing interests: None declared.

1 Mogensen CE, Neldam S, Tikkanen I, Oren S, Viskoper R, Watts RW, et al. Randomised controlled trial of dual blockade of renin-angiotensin system in patients with hypertension, microalbuminuria, and non-insulin dependent diabetes: the candesartan and lisinopril microalbuminuria (CALM) study [with commentary by A Berger]. *BMJ* 2000;321:1440-4. (9 December.)

Author's reply

Editor—McCormack's comments emphasise that more long term trials are needed in these patients. The reduction in urinary albumin:creatinine ratio did differ numerically between the combination treatment and lisinopril and candesartan, but the P value comparing treatment groups was not emphasised in the text. As we did not use the P value to claim that the specific comparison was significant it was not relevant to adjust for multiple testing.

It is common knowledge that there is always a small fall acutely in glomerular filtration rate or creatinine clearance with any antihypertensive treatment. This is usually associated with long term preservation of the glomerular filtration rate and is not harmful.

The comparison of baseline data with data at 24 weeks was not biased since there was only one initial randomisation (no new randomisation occurred after 12 weeks of single treatment). When the trial was designed there was concern that combination treatment with lisinopril and candesartan might cause hypotension (at that time thought to be a diastolic blood pressure of <80 mm Hg) since there was no experience then of combination treatment in this population.

As the UK prospective diabetes study 36 (UKPDS 36) has shown, a clear correlation is found between blood pressure and long term diabetic complications, with no J curve.[1] Any further reduction in blood pressure is extremely important and may suggest a new treatment strategy in type 2 diabetes.

Carl Erik Mogensen *professor of medicine*
Department of Medicine, M, Kommunehospitalet, University Hospital, DK-8000 Aarhus C, Denmark
cem@afdm.au.dk

Competing interests: CEM has received fees for speaking at symposiums supported by AstraZeneca, the manufacturer of lisinopril and candesartan cilexetil. He has also received funds for research and consulting fees from the company.

1 Adler AI, Stratton IM, Neil HAW, Yudkin JS, Matthews DR, Cull CA, et al. Association of systolic blood pressure with macrovascular and microvascular complications of type 2 diabetes (UKPDS 36): prospective observational study. *BMJ* 2000;321:412-9.

Air travel and risk of venous thromboembolism

Passengers should reduce consumption of alcohol on flights

Editor—Geroulakos,[1] like previous reviewers of the relation between air travel and venous thromboembolism,[2] did not mention the theoretical and experimental evidence of thrombogenesis in venous valve pockets that colleagues and I have published.[3] Modelled on one of the six possible permutations of Virchow's triad, our experiments produced experimental thrombi in venous valve pockets for the first time since Virchow described them in 1858.[4] The specific triad model was (1) interrupted circulation in venous valve pockets causing (2) hypoxaemic metabolic endothelial injury and leading to (3) ectopic haemostatic plug formation (blood metamorphosis) in valve pockets.

Merely to move blood clotting from position 1 to position 3 in the triad sequence gives a new explanation for thrombogenesis. This suggests that thrombogenesis during long haul flights is attributable to individual passengers' behaviour—specifically, taking an excess of drugs that suppress the central nervous system (alcohol, long acting tranquillisers, or other sedative drugs which, alone or in combination, may induce quasi-anaesthetic muscle paresis or paralysis).

During deep sleep, muscle areflexia may mean that muscles stop pumping blood towards the head and underperfuse deep venous valve pockets. The problem starts when non-pulsatile circulation into or within venous valves stops.[5] We did not establish the time for which valve pockets must be underperfused before their intima is suffocated and ectopic haemostatic thrombogenesis begins: our objective was to cause experimental thrombi, not prevent them. More than two hours' paralysis of limb muscles harmed valve pocket intima, but less than 90 minutes' paralysis produced no thrombi.

Geroulakos points out that airlines disclaim responsibility for thrombotic events and that no strict scientific basis exists for the standard medical advice given to passengers. Airlines may certainly disclaim responsibility if the lesions are caused by passengers' self injuring behaviour.

The standard advice to sober passengers is to move their legs, drink water, walk the aisle, be aware that there is a mild lack of oxygen, and seek more knee room if possible.

The advice to reduce consumption of alcohol and drugs that suppress the central nervous system is probably so pertinent that a legal limit on consumption on long haul flights might be introduced as prophylaxis against thrombotic disasters.

P C Malone *retired general practitioner*
129 Viceroy Close, Birmingham B5 7UY
pcmalone@Doctors.org.uk

1 Geroulakos G. The risk of venous thromboembolism from air travel. *BMJ* 2001;322:188. (27 January.)
2 House of Lords. *Deep vein thrombosis, seating, and stress.* London: Stationery Office, 2000: chapter 6. (Fifth report of select committee on science and technology.)
3 Hamer JD, Malone PC. Experimental deep venous thrombogenesis by a non-invasive method. *Ann R Coll Surg Engl* 1984;66:416-9.
4 Virchow R. *Cellular pathology.* New York: Dover, 1859:232 (chapter 10). [Translated by J Chance.]
5 Karino T, Motomiya M. Flow through a venous valve and its implications for thrombus formation. *Thromb Res* 1984;36:245-57.

Pulmonary embolism after air travel may occur by chance alone

Editor—In his editorial Geroulakos states that "there is only circumstantial, but no epidemiological, evidence connecting air travel with venous thrombosis."[1] He adds that the incidence of venous thrombosis associated with air travel is "much less than the impression given by the recent publicity" surrounding the death of a 27 year old woman from a pulmonary embolism after she disembarked from a flight from Australia to London.

This episode reminds me of a letter published 15 years ago in the *New England Journal of Medicine*.[2] Its authors reported a pulmonary embolism in a 40 year old man the day after he watched three consecutive football games on television on New Year's Day. The correspondents—presumably with tongue in cheek—termed this condition "bowl-game pulmonary embolism," in reference to the college football bowl games played in the United States in December and January.

I enjoy creative humour, but as an epidemiologist I felt compelled to point out that many pulmonary emboli will occur by chance alone among people viewing football games on television. In a letter (which the journal did not publish but which is on bmj.com[3]) I presented calculations estimating that pulmonary embolism would be expected to occur by chance alone in 34 "hard-core" viewers of bowl games during the 24 hours after the games.

One could do a similar calculation to estimate the expected occurrence of pulmonary embolism among the millions of people who have travelled by air during the past 24 hours. Some of those pulmonary emboli may be caused by the conditions of air travel that favour venous thromboembolism,[1] but many are probably related to air travel by mere coincidence.

I am not arguing against the sensible preventive measures that Geroulakos recommends at the end of his editorial. Rather, I am reinforcing his call for research to determine whether air travel is a genuine risk factor for venous thromboembolism, and to identify those at risk and the factors that correlate with risk.

Ronald M Davis *North American editor, BMJ*
Center for Health Promotion and Disease Prevention, Henry Ford Health System, Detroit, MI 48202-3450, USA
rdavis1@hfhs.org

1 Geroulakos G. The risk of venous thromboembolism from air travel. *BMJ* 2001;322:188. (27 January.)
2 Walsh C, Lauer MS, Huang PL, Page RL. Bowl-game pulmonary embolism. *N Engl J Med* 1986;314:1322-3.
3 Davis R. Pulmonary embolism occurs in some travelers by chance alone. Electronic response to: The risk of venous thromboembolism from air travel. bmj.com 2001;322. (bmj.com/cgi/eletters/322/7280/188#EL5; accessed 1 May 2001.)

Homoeopathy in malaria: head of infectious diseases replies

Editor—Delaunay et al report on a patient with complicated malaria after homoeopathic prophylactic treatment.[1] Besides pointing out the dangers of homoeopathic prophylaxis for this condition, the authors suggest that recommendations concerning good laboratory practice seem not to have been followed by the haematology laboratory.[2]

The patient was aware of the risk of contracting malaria, but her stay in the endemic zone was short (two weeks) and she had been told by her doctor that curative treatment would be available on her return should she develop the condition. Four days after her return she presented with fever, and she consulted her doctor 48 hours later. She was negative for *Plasmodium* spp at this stage. Her condition worsened, and she developed diarrhoea. A thick blood film two days later was again negative for plasmodia. After a further two days of parenteral antibiotic treatment, fever persisted at 40°C and she was admitted to a private clinic.

Blood examination showed anaemia, leucopenia, and thrombocytopenia, with hyperbilirubinaemia and moderately raised serum transaminase activity. She had dyspnoea, and extensive investigations were conducted (bronchial fibroscopy with bronchoalveolar lavage, abdominal and cardiological ultrasound examinations, blood cultures, serology tests). Her condition did not improve, and she was admitted to Nice University Hospital's intensive care unit.

There both thick and thin blood films showed the presence of *P falciparum*, with a 7% infection rate. She was treated with quinine according to existing guidelines, but her course was complicated by bacterial pneumonia, cholecystitis (requiring subsequent cholecystectomy), gastrointestinal bleeding, fungal systemic infection, and biological pancreatitis. She required assisted ventilation for 30 days, renal dialysis for six weeks, and multiple blood transfusions. A blood sample requested from the laboratory that performed the initial thick film was examined by the hospital parasitology laboratory and found to contain malarial parasites.

Several comments arise from this case. Firstly, because the level of parasitaemia in malaria varies from hour to hour, blood should have been examined several times a day for two or three days.[2] Secondly, the blood count and results of biological investigations are compatible with malaria. Thirdly, it is part of standard laboratory practice to investigate a low platelet count (in this case 66×10^9/l) further.[3]

This case shows the dramatic consequences of delayed diagnosis of malaria, both from the patient's standpoint and in view of the economic impact of a prolonged stay in hospital, extensive investigations, and serious medical and surgical complications.

Pierre Dellamonica *head, infectious and tropical diseases department*
Hôpital de l'Archet I, BP 3079, 06202 Nice cedex 3, France
mal-infectieuses@chu-nice.fr

1 Delaunay P, Cua E, Lucas P, Marty P. Homoeopathy may not be effective in preventing malaria. *BMJ* 2000;321:1288. (18 November.)
2 Goldsmith RS. Infectious diseases: protozoal. In: Krupp MA, Chatton MJ. *Current medical diagnosis and treatment.* Los Altos, CA: Lange Medical, 1983:878.
3 Ministère de Santé. *Journal officiel 1996. Législation et réglementation. Biologie médicale: nomenclature des actes.* Paris: Ministère de Santé, 1996.

Sifting the evidence

Likelihood ratios are alternatives to P values

Editor—In their critique of P values Sterne and Davey Smith omit two crucial reasons why P values do not adequately reflect evidence.[1]

Firstly, their statement (borrowed from Fisher) that "P values measure the strength of the evidence against the null hypothesis" does not stand up to scrutiny. A small P value means that what we observe is possible

but not very likely under the null hypothesis. But then life is made up of unlikely events. P values cannot deliver evidence against a hypothesis, no matter how low the cut-off point for saying that a result is significant. Short of P = 0, there is no such thing as evidence against a hypothesis.

Secondly, if evidence is what the data say then P values fail to qualify. P values are based on factors other than the observed data, notably on results "more extreme than these." The P value is literally the sum of probabilities of events that might have happened but did not. Furthermore, to compute a P value you must know what distribution to apply to those unobserved results.

Imagine a trial of vitamin C versus placebo in matched pairs of patients with the common cold; the number of pairs in which the patient taking vitamin C fares better is the outcome of interest. If the total number of observations was predetermined the P value is computed with the binomial distribution; if it was the smallest number of successes per group the negative binomial distribution applies.[2] The same trial result could lead to the null hypothesis being rejected or accepted depending on what you were told about the study design—that is, not on data alone. Other extraneous considerations that influence P values include the decision to use a one sided or a two sided test, and any adjustments made for multiple comparisons.

Several statisticians have proposed a solution: data cannot determine the absolute worth of one hypothesis taken in isolation but can provide evidence about the relative merit of two hypotheses specified a priori.[3–5] The data support hypothesis A over hypothesis B if the likelihood of the data is greater under A than B; the strength of evidence favouring A over B is the likelihood ratio. Not only is this approach compatible with logic but it considers only the observed data.

Thomas V Perneger *head*
Quality of Care Unit, Geneva University Hospitals, CH-1211 Geneva 14, Switzerland
Thomas.Perneger@hcuge.ch

1 Sterne JAC, Davey Smith G. Sifting the evidence—what's wrong with significance tests? *BMJ* 2001;322:226-31. (27 January.)
2 Berger JO, Berry DA. Statistical analysis and the illusion of objectivity. *Am Scientist* 1988;76:159-65.
3 Birnbaum A. On the foundations of statistical inference [with discussion]. *J Am Stat Assoc* 1962;53:259-326.
4 Hacking I. *Logic of statistical inference*. New York: Cambridge University Press, 1965.
5 Royall RM. *Statistical evidence. A likelihood paradigm*. London: Chapman and Hall, 1997.

Statistics must not be confused with science

EDITOR—I agree with much of what Sterne and Davey Smith say,[1] but the problems should be viewed more combatively. The reason why the editor of the *BMJ* thinks that doctors are deficient in statistics is because he confuses statistics with science.[2] This is why the *BMJ* is of more use to NHS managers than to researchers.

The confusion between a statistical hypothesis and a scientific one is widespread. A scientific theory does not have a distribution in the sense of probability theory. Newtonian physics is either right or wrong; there is not an infinite array of newtonian theories merging with those of Einstein. The two theories are qualitatively different, each ordering reality in a discontinuous way. They are not summaries of reality, nor do they have errors in a statistical sense. You don't do a systematic review of the Ptolemists and Copernicus and then do a Cochrane plot. The planets either move in a certain way or they don't.

This view of a scientific theory as something that brings coherence to nature, as a revealer of "hidden likenesses," has little to do with probability theory. The philosopher David Hume described the fatal weakness of inductionism: in a clinical context, the idea that you can take, say, 10 000 people with a stroke and then randomise them to either of two treatments and expect to get sense at the end is naive.

The argument about the importance of statistical power is minor. Sadly, the erroneous belief that small studies are unethical is now institutionalised by ethics committees. Should you study one hypothesis on 100 patients or test more hypotheses with smaller sample sizes? How to choose? Well, certainly not by performing power calculations. P values are not markers of truth. If they were we would have invented an epistemological engine. We haven't, nor can we, for reasons that Popper laid out formally but really are obvious. Or do we imagine that those busy systematic souls who go around adding up other people's P values will now do systematic reviews of quantum mechanics, linguistics, etc and reveal the structure of nature. No, of course not.

The idea that averaging the data is a way to understand nature would be laughable if it didn't do so much harm to genuine clinical discovery.

Jonathan Rees *professor of dermatology*
University of Edinburgh, Edinburgh EH8 9AG
jonathan.rees@ed.ac.uk

1 Sterne JAC, Davey Smith G. Sifting the evidence—what's wrong with significance tests? *BMJ* 2001;322:226-31. (27 January.)
2 Editor's choice. Some gentle statistics. *BMJ* 2001;322 (7280). (27 January.)

Perfect understanding seldom happens

EDITOR—The 27 January issue of the *BMJ* is particularly thought provoking (see for a start Editor's Choice[1]). Sterne and Davey Smith illustrate the fallibility of tests of significance.[2] Socially responsible people are shocked that the government promotes a dogma and then finds evidence to back it.[3] It seems that a guideline may be considered good, bad, irrelevant, wrong, premature, or tardy depending on who is speaking.[4]

Colleagues: let us mistrust everything. Whereas once I thought this a cynical cop-out, I now realise that it is an intellectually respectable stance, meeting Sterne and Davey Smith's explanation of a bayesian position on statistical truth. If I understand them correctly, this means: "This is what I think I know. Now let's see if you can shake my view."

Given that the validity and statistical robustness of evidence is so fragile, what else are practitioners to do? Certainly we should not put our trust in consensus statements, such as those emanating from the National Institute for Clinical Excellence. Although these statements may be an improvement on "tendentious opinions selectively heralded" (TOSH), they may prove to be nothing more than "current right advice … probably" (CRAP).

Such edicts are relics of the time when it was a defence against complaint to appeal to an agreed body of professional opinion. With evidence as contentious as that highlighted by articles in the journal, the test of "best current opinion" becomes illusory. There will always be another expert opinion a standard deviation away. Ironically, as evidence becomes devalued and more relativistic, arriving at a considered judgment becomes more important. Such judgment used to be called professionalism; it predated guidelines until it was undermined by the joint efforts of the General Medical Council and the Department of Health.

In my view, a new institute is needed to support those of us hoping to preserve professional medical practice, in which we try to do our best for our patients, taking into account their individual circumstances and wishes while drawing on our experience of practising medicine. The institute will not deny the need for research or the importance of the P value. It will embrace controlled trials yet not dismiss n = 1 studies. It will accept guidelines but only as aide memoires, not as gospel. Everyone else has an institute with a heart warming title, so I propose that this one is called the institute for "perfect understanding seldom happens; opinion flirts with facts." PUSH OFF will do nicely as an acronym.

Michael Apple *general practitioner*
Garston Medical Centre, Watford WD25 9GP
m.apple@virgin.net

1 Editor's choice. Some gentle statistics. *BMJ* 2001;322 (7280). (27 January.)
2 Sterne JAC, Davey Smith G. Sifting the evidence—what's wrong with significance tests? *BMJ* 2001;322:226-31. (27 January.)
3 Macintyre S, Chalmers I, Horton R, Smith R. Using evidence to inform health policy. *BMJ* 2001;322:222-5. (27 January.)
4 Lux AL, Edwards SW, Osborne JP, Hancock E, Johnson AL, Kennedy CR, et al. Revised guideline for prescribing vigabatrin in children. *BMJ* 2001;322:236-7. (27 January.)