considered, it was unnecessary as the areas of ulceration healed completely after resolution of the infection.

I agree that cutaneous radionecrosis and osteonecrosis often require excision and the use of myocutaneous flaps. These and other complications of radiotherapy are covered in a subsequent article in the ABC series. Reconstruction of the breast and its complications are also being covered separately. The point of the illustration was to show a good example of cellulitis after surgery and radiotherapy. This cellulitis needs to be controlled by antibiotics before other measures are considered. In this respect I do not believe that the illustration was misleading or inappropriate.

J MICHAEL DIXON
Consultant surgeon

Edinburgh Breast Unit,
Western General Hospital,
Edinburgh EH4 2XU

Macroglossia

Classify the pathology as normal or abnormal

EDITOR,—We agree with P Murthy and M R Laing that the term macroglossia should be restricted to chronic enlargement of the tongue as the differentiation from acute swelling has implications for both the aetiology and management.[1] Having made this distinction early in the paper, the authors later give rise to confusion by listing angio-oedema, Ludwig's angina, and angioneurotic oedema as causes. These conditions are of rapid onset and do not cause true macroglossia. We have reported on a series of patients with acute swelling of the tongue and found corticosteroids useful.[2] They are unlikely, however, to be of benefit in long standing cases.

In our experience, it is helpful to classify the pathology of macroglossia as normal or abnormal. Biopsy is therefore indicated, particularly if reduction glossectomy is contemplated. The diseased tongue—for instance, in lymphangioma or in amyloidosis—is a hugely vascular structure, and blood loss may be considerable and even life threatening. Possible preoperative precautions include tracheotomy, ligation of the external carotid artery and embolisation of selective branches of the external carotid artery tree. Surgery by laser is a theoretical alternative, but experience is limited.

A G RENEHAN
Medical student

University of Manchester,
Manchester M13 9PL

M E MORTON
Consultant in oral and maxillofacial surgery

Blackburn Royal Infirmary,
Blackburn BB2 3LR

1 Murthy P, Laing MR. Macroglossia. *BMJ* 1994;**309**:1386-7. (26 November.)
2 Renehan A, Morton M. Acute enlargement of the tongue. *Br J Oral Maxillofac Surg* 1993;**31**:321-4.

Tracheostomy may not be necessary during surgery

EDITOR,—The literature does not usually support P Murthy and M R Laing's view that a tracheostomy is required to maintain the airway during surgery for macroglossia.[1] Two more conservative alternatives exist. Nasal intubation in preparation for surgery is the optimal anaesthetic approach. The airway is then maintained in safety[2,3] with the endotracheal tube in position for several days after surgery while the child is monitored in a paediatric intensive care unit. This allows the postoperative swelling of the tongue to recede. The other, slightly unconventional method is to secure the tongue with a 2-0 silk suture placed in the tip of the tongue and then tape the suture to the cheek.[4]

Though the latter method may well have a role, the former is logical as any upper airway obstruction due to oedema of the tongue will be only temporary. This approach obviates the need for a tracheostomy at the time of the initial surgery, maintains a safe airway, and allows for the option of a tracheostomy if the oropharyngeal airway is not deemed to be safe at a trial of extubation.

R W R FARRELL
Registrar in otolaryngology

Department of Otolaryngology,
Doncaster Royal Infirmary,
Doncaster DN2 5LT

1 Murthy P, Laing MR. Macroglossia. *BMJ* 1994;**309**:1386-7. (26 November.)
2 Kveim M, Fisher JC, Jones KL, Gruer B. Early tongue resection for Beckwith-Wiedemann macroglossia. *Ann Plast Surg* 1985; **14**:142-4.
3 Sokoloski PM, Ogle RG, Waite DE. Surgical correction of macroglossia in Beckwith-Wiedemann syndrome. *Journal of Oral Surgery* 1978;**36**:212-5.
4 Patterson GT, Ramastry SS, Davis JU. Macroglossia and ankyloglossia in Beckwith-Wiedemann syndrome. *Oral Surg Oral Med Oral Pathol* 1988;**65**:29-31.

Confounding and Simpson's paradox

Multiple regression would confound the clinicians

EDITOR,—Steven A Julious and Mark A Mullee discuss confounding and Simpson's paradox, defining confounding as occurring "when the association between an exposure and an outcome is investigated but the exposure and outcome are strongly associated with a third variable. An extreme example of this is Simpson's paradox, in which this third factor reverses the effect first observed."[1] In their analysis of my and colleagues' article on treatments for surgery to remove kidney stones the authors correctly observe that the size of the stones is the single most important factor in determining the success of treatment. The fact that the figures show that open stone surgery has a higher success rate for both small and large stones but that the overall success rate is greater with percutaneous nephrolithomy is a paradox. This probably reflects the numbers in each group.

The authors' deduction that factors such as the patient's age and characteristics determine treatment does not, however, hold for a historical perspective, when all patients were treated consecutively with the only treatment available at this time. There is no reason to believe that patients treated 10 or 15 years ago had any different characteristics from those of patients presenting for modern day treatments.

The authors' argument that randomised trials are necessary to show any effect of treatment is difficult to accept when you have a new form of treatment that is so clearly superior to all previous forms of treatment that to compare one against the other—that is, open surgery versus lithotripsy—for a small stone would be unethical with regard to morbidity.

The main aim of the paper was not really to compare success rates but more to show that, when success rates were comparable, morbidity and cost to the health service were much reduced. I therefore do not see what purpose could be served by producing a model to include multiple regression or multiple logistic regression in analysing the variance as it would only confound the clinicians.

C R CHARIG
Consultant urologist

Epsom General Hospital,
Epsom,
Surrey KT18 7SG

1 Julious SA, Mullee MA. Confounding and Simpson's paradox. *BMJ* 1994;**309**:1480-1. (3 December.)

Authors' reply

EDITOR,—C R Charig states that the reason why open surgery has a higher success rate for both small and large stones but a poorer success rate overall when compared with percutaneous nephrolithotomy is the numbers in each group. This, however, does not seem to be the case. Substantially more patients having open surgey had large stones, and as the size of the stones is the most important factor in determining success rates it was probably this difference that gave open surgery a poorer overall success rate.

There is reason to believe that patients' characteristics may have changed over time. Altogether 75% (263/350) of patients undergoing open surgery between 1972 and 1980 had large stones, compared with only 23% (80/350) undergoing percutaneous nephrolithotomy between 1980 and 1985. This seems to indicate a change in the ratio of large to small stones treated.

If advances in treatment are to be made it is necessary to establish that "a new form of treatment is so clearly superior to all previous forms of treatment." Sheldon stated, "in non-randomised observational studies..., patients receiving different treatments may differ systematically with respect to any number of known and unknown factors that affect prognosis.... Although statistical adjustments may be made in an attempt to exclude the effects of these confounders..., this assumes both a complete knowledge of the confounding variables and their comprehensive and accurate measurement. Neither is likely to be possible and at least a moderate bias will remain."[1] To avoid bias a randomised controlled trial should be performed to ensure that the new treatment is truly superior. The perils of avoiding or ignoring randomised controlled trials have been further described (G D Smith and T A Sheldon, meeting of the Society of Social Medicine, Cambridge, September 1993). After favourable reports in the 1960s clofibrate came to be widely prescribed and various descriptive and non-randomised studies that purported to show its efficacy were published. After unfavourable results of two large randomised controlled trials were reported, however, clofibrate was rarely used.

It is unethical not to use a new treatment that is clearly superior, but "is it not unethical to advise patients to have an operation that has not been formally tried and tested?"[2]

Statistical methods, such as analysis of covariance, give the same measure of effect as more simple analyses but have the advantage of having been adjusted for possible imbalances between groups.

MARK A MULLEE
Senior statistical programmer
STEVEN A JULIOUS
Statistician

Medical Statistics and Computing,
University of Southampton,
Southampton General Hospital,
Southampton SO16 6YD

1 Sheldon TA. Please bypass the PORT: observational studies run a poor second to randomised controlled trials. *BMJ* 1994;**309**: 142-3.
2 Surgical innovation under scrutiny [editorial]. *Lancet* 1993;**342**: 187-8.

Matching in case-control studies

EDITOR,—J Martin Bland and Douglas G Altman's review of the advantages and disadvantages of matching in case-control studies omits two important problems.[1]

Firstly, matching in case-control studies ensures that the matching factors, such as age or sex, are equally distributed between cases and controls. Although matching thus removes the original

confounding effect of these factors, it may introduce a new bias. If the matching factor is associated with the exposure of interest a crude case-control analysis may lead to a result that is confounded by the matching factor itself.[2] In such a case the confounding must be eliminated by stratified analysis or multivariate analysis.

Another potential problem is overmatching, which results from cases and controls being matched on variables that are highly related to the exposure. For example, suppose one matches on the variable "carrying a cigarette lighter" in a study of smoking and lung cancer. Matched analyses depend on the existence of discordant pairs of cases and controls, but few will exist in this situation because the confounder is closely related to the exposure. Overmatching will not bias the results, but it will make the analysis highly inefficient, which abrogates the main virtue of matching.[2]

In case-control studies investigators must use the technique of matching wisely since it is irreversible and its advantages may be limited by the problems we have highlighted.

HENRIK TOFT SØRENSEN
Associate professor

Danish Epidemiology Science Centre,
University of Aarhus,
8000 Aarhus C,
Denmark

MATTHEW W GILLMAN
Assistant professor

Department of Ambulatory Care and Prevention,
Harvard Medical School and Harvard Community Health Plan,
126 Brookline Avenue,
Boston, MA 02215, USA

1 Bland JM, Altman DG. Matching. *BMJ* 1994;**309**:1128. (29 October.)
2 Rothman K. *Modern epidemiology*. Boston: Little, Brown, 1986.

Numbers alone cannot determine rational treatment

EDITOR,—In the Editor's Choice in the issue of 12 November meta-analysis is embraced with enthusiasm as a means of determining effective treatment and solving therapeutic dilemmas. This approach should form part of the armamentarium of those interested in rational prescribing and seems to ensure that data from patients who have given up time and exposed themselves to risk by participating in clinical studies are used to full advantage. But how reliable are meta-analyses?

Recently it has been possible to compare the results of meta-analysis with those of a prospective study. For example, the fourth international study of infarct survival included over 50 000 patients, substantially more than the 1000-4000 combined to form meta-analyses of the same therapeutic area.[1,2] In many respects the results of the fourth international study of infarct survival and the meta-analyses are similar, but in others they differ: whereas meta-analysis showed clear benefit for intravenous magnesium and nitrates in treating acute myocardial infarction, the study of infarct survival did not. It is unlikely that the meta-analyses were "wrong" because of statistical problems as several groups came independently to the same conclusion. Furthermore, owing to the numbers involved, any subsequent analysis that included the fourth international study of infarct survival would reach the same conclusion as that study. A meta-analysis seems no more infallible than the clinical trials from which it is composed.

Emphasis has been placed on ensuring that only trials with the correct "statistical" design are included in meta-analyses, but less interest has been shown in "biological" design. The case of magnesium again illustrates the problem. In the second Leicester intravenous magnesium intervention trial, a large study of magnesium in acute myocardial infarction, the intravenous magnesium was given early and before thrombolysis, and benefit was seen.[3] In the fourth international study of infarct survival magnesium was given later and after thrombolysis, and no benefit was seen. Analysis based on numbers alone would suggest that magnesium has no advantage. However, analysis that takes into account the timing of administration suggests that early intervention might be beneficial,[4] a conclusion consistent with data from studies in animals indicating prevention of reperfusion "injury" by magnesium.[4]

Decisions on treatment should be based on best available evidence, and this often falls short of certainty. However, in the justifiable rush to establish evidence based treatment large numbers should not swamp critical clinical assessments. Meta-analyses and large trials are useful tools and have helped to shape rational treatment protocols, but greater account of therapeutic considerations should be taken when including or excluding trials. Prescribers should be prepared for frequent changes in treatment protocols as best available evidence is continually updated.

PATRICK VALLANCE
Senior lecturer in clinical pharmacology

Clinical Pharmacology Unit,
Department of Pharmacology and Clinical Pharmacology,
St George's Hospital Medical School,
London SW17 0RE

1 Horner SM. Efficacy of intravenous magnesium in acute myocardial infarction in reducing arrhythmias and mortality. Meta-analysis of magnesium in acute myocardial infarction. *Circulation*. 1992;**86**:774-9.
2 Teo KK, Yusuf S. Role of magnesium in reducing mortality in acute myocardial infarction. *Drugs* 1993;**46**:347-59.
3 Woods KL, Fletcher S. Long-term outcome after intravenous magnesium sulphate in suspected acute myocardial infarction: the second Leicester intravenous magnesium intervention trial (LIMIT-2). *Lancet* 1994;**343**:816-9.
4 Casscells W. Magnesium and myocardial infarction. *Lancet* 1994;**309**:807-8.

Apolipoprotein e4 allele and cognitive decline

May be less relevant

EDITOR,—In their paper on the association of the e4 allele of apolipoprotein E and incidence of cognitive impairment in elderly men Edith Feskens and colleagues claim that 22% of incident cases of cognitive impairment can be attributed to the effect of the e4 allele.[1] This contrasts with much higher figures previously reported in patients with Alzheimer's disease,[2,3] but is in accordance with data that we collected on a series of patients with the disease.

Attributable fraction is the proportion of cases that would be avoided by reducing risk to the level found in the lower risk group, in this case subjects not carrying the e4 allele. Attributable fraction has been reported as ranging from 78% in familial Alzheimer's disease[3] to 53% in the sporadic disease,[2] indicating that a relevant proportion of cases of Alzheimer's disease might be due to the e4 allele and that other risk factors might have a minimal role. Attributable fraction was estimated, however, on the basis of the risk computed on prevalent cases (odds ratio)—that is, the risk of having the disease, assuming that disease duration was similar across e4 genotypes. In fact, prevalence, incidence, and disease duration are related in subjects homozygous for e4 (e4/e4) and in those not carrying the e4 allele (-/-) as follows:

$$\frac{P_{e4/e4}}{P_{-/-}} \quad \frac{I_{e4/e4}}{I_{-/-}} \quad \frac{D_{e4/e4}}{D_{-/-}}$$

$$\text{that is, odds ratio} = \text{relative risk} \times \frac{D_{e4/e4}}{D_{-/-}}$$

where P, I, and D are prevalence, incidence, and disease duration in patients with Alzheimer's disease for the relative e4 genotype. A corresponding relation applies to subjects heterozygous for the e4 allele. Longer disease duration in subjects with Alzheimer's disease carrying the e4 allele might lead to an overrepresentation of e4 in prevalent cases, thus resulting in odds ratios overestimating relative risk—that is, the risk of developing the disease. In this case, the relative risk, if computed on the basis of an inflated odds ratio, gives an inflated estimate of the attributable fraction. Even if as many as half of the cases of cognitive impairment are due to Alzheimer's disease and the e4 allele is not associated with other forms of dementia (which, however, does not seem to be true)[4,5] the proportion of cases of Alzheimer's disease attributable to the e4 allele according to Feskens and colleagues should not exceed 44%. This is in accordance with our own data on a hospital based series of 62 subjects with sporadic Alzheimer's disease that began at age 70 or over, which showed that the attributable fraction computed without taking longer disease duration in e4 carriers into account was 51% and 37% after correction for disease duration.[5]

Our and Feskens and colleagues' data show that the epidemiological relevance of the e4 allele in determining cognitive impairment might be lower than previously suggested, and that large, population based epidemiological studies need to be carried out to evaluate the risk of developing dementia in the many people who carry the e4 allele.

G B FRISONI
Staff specialist
A BIANCHETTI
Head

Alzheimer's Disease Unit,
Ospedale S Cuore-FBF,
I-25125 Brescia,
Italy

G FRANCESCHINI
Professor of pharmacology

Centre E Grossi Paoletti,
Institute of Pharmacological Sciences,
University of Milan,
I-20162 Milan

S GOVONI
Professor of pharmacology

Institute of Pharmacological Sciences,
University of Milan,
I-20133 Milan

M TRABUCCHI
Profesor of neuropharmacology

Geriatric Research Group,
I-25122 Brescia

1 Feskens EJM, Havekes LM, Kalmijn S, de Knijff P, Launer LJ, Kromhout D. Apolipoprotein e4 allele and cognitive decline in elderly men. *BMJ* 1994;**309**:1202-6. (5 November.)
2 Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC Jr, *et al.* Protective effect of apolipoprotein e type 2 allele for late onset Alzheimer's disease. *Nat Genet* 1994;7:180-4.
3 Payami H, Kaye J, Heston LL, Bird TD, Schellenberg GD. Apolipoprotein E genotype and Alzheimer's disase [letter]. *Lancet* 1993;**342**:738.
4 Frisoni GB, Bianchetti A, Govoni S, Trabucchi M, Calabresi L, Franceschini G. Association of apolipoprotein E e4 with vascular dementia [letter]. *JAMA* 1994;271:1317.
5 Frisoni GB, Govoni S, Trabucchi M, Francéschini G. Reply to letters by A Roses *et al* and T Pirttilä *et al* [letter]. *JAMA* (in press).

Author's reply

EDITOR,—Firstly, we agree with Frisoni and colleagues that to calculate the attributable fraction the use of prospective data (incidence of disease) is much more informative than the use of cross sectional data (prevalence of disease). The duration of the disease, due to earlier onset as well as selective survival, is likely to affect the results. Population based prospective studies on the occurrence of Alzheimer's disease, however, are difficult to accomplish because of the large baseline population needed.

Secondly, on the basis of the prevalence data in our study the attributable fraction can be estimated to be about 10%. This is clearly lower than that calculated from our longitudinal data, in contrast with the hypothesis of Frisoni and colleagues. This is probably mainly a numerical issue: comparing prevalence rates of 31% (men without e4) and 41% results in a lower odds ratio than comparing incidence rates of 16% and 28%, although the