maintain operation of the system. When a system is upgraded it provides additional features that require the original cost-benefit analysis to be reassessed. Upgrading the equipment commonly results in faster throughput, better discrimination, and added diagnostic value—for example, that provided by angiography—with the result that it tends to be cost neutral on a per patient basis.

In their analysis Szczepura and colleagues calculated that changes in patients' management helped to reduce the costs of examination by £80 and, furthermore, that if half of the diagnostic procedures performed in addition to magnetic resonance imaging were dropped this could result in an additional saving of £62.

The table shows that if the equipment is operated efficiently during the normal working day and with the above assumptions magnetic resonance imaging can make a positive contribution to health care costs. More importantly, if the service is operated for longer each day—a policy that is becoming more widely adopted by units in the United Kingdom—this contribution can be considerable.

Szczepura and colleagues' study emphasises that equipment with high fixed costs can be justified only if it is to be operated efficiently. By assessing and quantifying the outcome in addition to the costs the authors have established a more rational basis on which decisions to purchase magnetic resonance imaging equipment can be made, which I believe should be more widely adopted.

ANTHONY STEVENS

London SW6 4RF

1 Szczepura AK, Fletcher J, Fitz-Patrick JD. Cost effectiveness of magnetic resonance imaging in the neurosciences. *BMJ* 1991; **303**:1435-9. (7 December.)

Spinal manipulation and mobilisation for back and neck pain

SIR,—In B W Koes and colleagues' review of two decades of reports of trials of spinal manipulation the methods score used seems not to have been sufficiently adapted to the special requirements of such trials.[1] This deficiency may best be seen by considering its application to one of the trials reviewed.

Hadler *et al* stated that patients were randomly allocated,[2] but in Koes and colleagues' paper they lost all four available points for randomisation by not having stated the method. All 12 points available for adequacy of group sizes were lost as a figure of 50 patients per group was not reached. But the trial showed a highly significant effect of treatment ($p=0{\cdot}009$). Target group sizes in protocols are merely estimates of the numbers that will be needed to show an effect of treatment; once a positive result of high significance has been obtained the estimate is superseded by reality. To penalise this trial for a demonstrably adequate group size is illogical.

The trial is not awarded the five points available for use of a placebo control as the sham manipulation employed involved laying on of hands and may thus have had some physical beneficial effect. If this had been so it would have led to an underestimate of the benefit from manipulation and therefore could not invalidate a positive result.

Ten points were available for using five different outcome measures, and yet Pocock *et al*, whom Koes and colleagues quote, advise deciding a priori on a small number of outcome measures and end points to avoid invalidating the significance tests used.[3] Other measures may be made as an exploratory feature of the design to compare the utility of different outcome measures, but this secondary function provides data for use in designing further trials and is not relevant when the trial's primary function of assessing outcome is being considered. Credit in this section should be given for the authors stating prospectively a small number of appropriate outcome measures. Hadler *et al* used one: the disability score designed by Roland and Morris and shown by them to be a more reliable and sensitive index of disability in back pain than measures such as pain experienced or spinal mobility,[4] for which Koes and colleagues would have awarded points.

Ten points were available if the five suggested outcome measures were measured blind. Hadler *et al* relied on a patient questionnaire administered over the telephone by someone unaware of the treatment, and their control patients had received the most realistic "sham" treatment of all reported trials. Therefore blind assessment was probably optimum and yet no points were awarded.

If the same categories and weighting that Koes and colleagues used were applied, the adjustments suggested above would increase the score for Hadler *et al*'s trial from 53% to 90% of the maximum. This confirms my view that this trial had the most sophisticated design reported to date and was greatly undervalued in the review.

R S MACDONALD

London College of Osteopathic Medicine,
London NW1 6QH

1 Koes BW, Assendelft WJJ, van der Heijden GJMG, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain. *BMJ* 1991;**303**:1298-303. (23 November.)
2 Hadler NM, Curtis P, Gillings DB, Stinnett S. A benefit of spinal manipulation as adjunctive therapy for acute low-back pain: a stratified controlled trial. *Spine* 1987;**12**:703-6.
3 Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. *N Engl J Med* 1987;**317**:426-32.
4 Roland M, Morris R. A study of the natural history of back pain. Part 1: development of a reliable and sensitive measure of disability in low-back pain. *Spine* 1983;**8**:141-4.

SIR,—B W Koes and colleagues' review of trials of manipulation for back pain has two serious limitations.[1] Firstly, many of the criteria and methods are arbitrary and illogical. A less than homogeneous study population reflects the subjects seen in real life and may increase the applicability of results, particularly where (as in our case) minimisation is used to permit analyses for different groupings, an advantage that Koes and colleagues did not recognise.

It is not mandatory to avoid "cointerventions" (other treatments) in a pragmatic trial, as resort to other treatments may in day to day practice be the consequence of the approaches under comparison. Indeed, insistence on pragmatic trials and the simultaneous avoidance of cointerventions, which is what Koes and colleagues imply by their criteria, make it impossible to recognise the full implications of different policies. These inconsistent and debatable criteria seem to carry as much weight in the scoring system as trial size. Many trials in back pain set out to compare different active regimens in circumstances in which placebo treatment would be unethical, so that penalising these trials for the absence of a group treated with placebo is inappropriate. There are obvious difficulties in blinding patients in trials of different manipulative techniques.

Secondly, there are several inaccuracies. Our trial of chiropractic and hospital management for back pain—not specifically of manipulation—came near the top of the scores derived, but we refer to it here only to exemplify these errors.[2] There was a surprisingly high proportion (20%) of initial mistakes by the reviewers "usually . . . due to errors in reading," a problem already identified.[3] In the case of our own trial the authors did not recognise that we did establish homogeneous subgroups or that we described interventions in both the text and a table. Even more obviously, the authors failed to acknowledge our explicit statement that we analysed the results by intention to treat. One trial[4] was included in this and a previous review by Koes *et al*[5] because of its factorial design, a considerable advantage overlooked by the authors. Although almost identical criteria were used, it scored 38 on one occasion and 50 on the other, suggesting a serious degree of inconsistency in the review process.

Whether the term manipulation should be used, as Koes and colleagues used it, "to cover both manipulation and mobilisation" is at least arguable. Their conclusions about manipulation (thus defined) in back pain seem analogous to generalisations about antibiotics for sore throat.

T W MEADE
JOY TOWNSEND
ANDREW FRANK

MRC Epidemiology and Medical Care Unit,
Northwick Park Hospital,
Harrow,
Middlesex HA1 3UJ

1 Koes BW, Assendelft WJJ, van der Heijden GJMG, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ* 1991;**303**:1298-303. (23 November.)
2 Meade TW, Dyer S, Browne W, Townsend J, Frank AO. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *BMJ* 1990;**300**:1431-7.
3 Meade TW. Effectiviteit van chiropractie en fysiotherapie bij behandeling van lage rugpijn. *Nederlands Tijdschrift voor Manuele Therapie* 1991;**10**:14-6.
4 Coxhead CE, Inskip H, Meade TW, North WRS, Troup JDG. Multicentre trial of physiotherapy in the management of sciatic symptoms. *Lancet* 1981;**i**:1065-8.
5 Koes BW, Bouter LM, Beckerman H, van der Heijden GJMG, Knipschild PG. Physiotherapy exercises and back pain: a blinded review. *BMJ* 1991;**302**:1572-6. (29 June.)

SIR,—The blinded review of spinal manipulation and mobilisation by B W Koes and colleagues, from the department of epidemiology and biostatistics at the University of Limburg, is welcome for the attention it draws to a difficult problem and the meticulous scientific methods used.[1] The absence of clinical input, however, has led to erroneous judgments.

I wonder, for example, in respect of our own paper on what basis the authors gave it a score of 0 out of 5 for "mentioning good qualifications of manipulative therapist."[2] We clearly were too modest in assuming that membership of the Chartered Society of Physiotherapists and working and teaching in the department in which Cyriax pioneered the treatment under study were adequate. With similar modesty I have rescored our study and find that it now scores 90 and is top of the list. This is not a surprise as the study was mounted with a careful design and after advice from both clinicians and statisticians.

The serious message is that epidemiologists and statisticians may not be qualified to assess the merits of clinical papers and that their pronouncements may be misleading. I believe that such assessments should use appropriate clinicians.

J A MATHEWS

Department of Rheumatology,
St Thomas's Hospital,
London SE1 7EH

1 Koes BW, Assendelft WJJ, van der Heijden GJMG, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ* 1991;**303**:1298-303. (23 November.)
2 Mathews JA, Mills SB, Jenkins VM, Grimes SM, Morkel MJ, Mathews W, *et al*. Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *Br J Rheumatol* 1987;**26**:416-23.

SIR,—In their review of the most worthy papers on manipulation for back and neck pain Koes and colleagues have shown the unsatisfactory nature of almost all previous work.[1] We do not, however, agree with the suggestion that further attempts should be made to answer the same question by using this format. Of the papers reviewed by the authors, four achieved methodological scores of 50-60, and these were all comparatively recent. We think it unlikely that the quality of this type of study could improve dramatically, and this is

supported by a report by the Department of Health and Social Security's working group on back pain, which analysed many papers on this subject in terms of scientific validity and found only three on manipulative therapy that fulfilled its criteria.[2] In general these studies are akin to asking "Is appendicectomy an effective treatment for abdominal pain?"

Breen, a chiropractor, suggested that "for those who use manipulative therapy in the management of back pain patients, a clear idea of the nature of pain and its possible sources is of paramount importance."[3] Furthermore, "without rational hypotheses to address these questions treatment becomes incoherent and irrelevant to the problem." By way of example, mechanical and local inflammatory sources of pain may be amenable to manual therapy, but infective, metabolic, neoplastic, and systemic inflammatory causes most probably are not. Indeed, it is quite possible that acute and chronic conditions, whether in the neck or in the back, will all respond in different ways. In keeping with Breen's view we believe strongly that any future attempt to rationalise the role of manipulative therapy in the treatment of back pain should be preceded by an attempt to classify and segregate both site and cause.

S M HAY
B TODD

Department of Orthopaedic Surgery,
Northern General Hospital,
Sheffield S5 7AU

1 Koes BW, Assendelft WJJ, van der Heijden GJMG, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ* 1991;**303**:1298-303. (23 November.)
2 Department of Health and Social Security. *Working group on back pain.* London: HMSO, 1979.
3 Breen A. Sources of back pain due to mechanical disorders within the lumbar spine: a chiropractor's view. *Proceedings of the Society for Back Pain Research* 1990 May:9-11.

AUTHORS' REPLY,—We are pleased that three of the letters discussing our review are written by authors of trials that we listed in our "top 10."[1] The detailed comments on the methods we used to assess published trials show the considerable advantage of this method of review. Whereas traditionally authors of review articles implicitly apply their (unknown) criteria, we explicitly formulated and used our criteria, which obviously facilitates discussion.

R S MacDonald states that the trial of Hadler *et al*[2] was greatly undervalued in the review, but this trial scored 53 points, which was the second best score. MacDonald applied our criteria and weighting to Hadler *et al*'s trial and ended up with 90 points. J A Mathews assessed a study of which he was a coauthor[3] and also obtained 90 points. Both of these scores seem to be based on a different use of our criteria and the use of additional information that was not included in the original articles.

The statement that random allocation has been carried out is, in our opinion, not sufficient to warrant all four points for this criterion. All studies included in our review were randomised controlled trials. Studies could earn points if there was a clear description of how the randomisation procedure had been carried out. For readers this is important.[4] The points given for sample size are not meant as a reward for sufficient power. Consequently we do not agree that points should be given depending on whether a certain difference in outcome was significant. Our main reason for giving points for large sample sizes was the assumption that with larger numbers one has more assurance that (un)known prognostic indicators will be equally divided between the study groups.

MacDonald's suggestion that all 10 points should be given for relevant outcome measures when an author reports only one or two primary outcome measures is certainly interesting. Of course these primary outcome measures should be chosen (and preferably published) before the data from the trial are analysed. Otherwise the reader has no assurance that the choice of primary outcome measures was not guided by the data analysis.

T W Meade and colleagues performed one of the best studies,[5] but it still showed severe shortcomings.[6] They think that we insisted on pragmatic trials, but this is obviously not the case because studies were also rewarded if they included an adequate placebo treatment. Meade and colleagues claim points for having carried out an analysis based on intention to treat. In our review, however, studies could earn points for this criterion only if in cases of more than 10% loss to follow up (which was the case after two years' follow up) and an alternative analysis had also been carried out correcting for withdrawals and missing values.

We do not agree with Mathews that these kind of reviews can be carried out only by clinicians (although three of us are clinicians). As long as explicit criteria are used that can be applied and checked by most readers we see no ground for this statement.

We agree with S M Hay and B Todd that the design of studies of this treatment is probably difficult to improve. A recent study from our department scored 55 points when assessed by the two reviewers (GJMGH and WJJA, in this case not blinded) who had made the assessments in our meta-analysis.[7] We still think, however, that further trials are needed to determine the efficacy of spinal manipulation and mobilisation for well defined subgroups of patients with back and neck pain. Although the methods of reviewing that we used clearly need to be developed further, we recommend them for reviewing past and future studies.

BART KOES GEERT VAN DER HEIJDEN
PIM ASSENDELFT LEX BOUTER
PAUL KNIPSCHILD

Department of Epidemiology and Biostatistics,
University of Limburg,
PO Box 616,
6200 MD Maastricht,
The Netherlands

1 Koes BW, Assendelft WJJ, van der Heijden GJMG, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ* 1991;**303**:1298-303. (23 November.)
2 Hadler NM, Curtis P, Gillings DB, Stinnett S. A benefit of spinal manipulation as adjunctive therapy for acute low-back pain: a stratified controlled trial. *Spine* 1987;**12**:703-6.
3 Mathews JA, Mills SB, Jenkins VM, Grimes SM, Morkel MJ, Mathews W, *et al*. Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *Br J Rheumatol* 1987;**26**:16-23.
4 Altman DG. Randomisation: essential for reducing bias. *BMJ* 1991;**302**:1481-2.
5 Meade TW, Dyer S, Browne W, Towsend J, Frank AO. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *BMJ* 1990;**300**:1431-7.
6 Assendelft WJJ, Bouter LM, Kessels AGH. Effectiveness of chiropractic and physiotherapy in the treatment of low back pain—a critical discussion of the British randomized clinical trial. *J Manipulative Physiol Ther* 1991;**14**:281-6.
7 Koes BW, Bouter LM, Mameren van H, Essers AHM, Verstegen GMJR, Hofhuizen DM, *et al*. Randomised clinical trial of manual therapy and physiotherapy for persistent back and neck complaints: results of one year follow up. *BMJ* (in press).

Glasgow coma scale and gag reflex

SIR,—With regard to A J Stanners's comments[1] on our paper on the relation between the Glasgow coma scale and the gag reflex,[2] the importance of other reflexes and mechanisms in protecting the airway is beyond dispute. The paper that we originally submitted to the *BMJ* discussed these more fully and included data on the cough reflex in the patients, which showed that the cough reflex was much less easily suppressed than the gag reflex.

The cough and swallow reflexes and other protective mechanisms are not, however, easy to assess objectively, especially in an accident and emergency department, where sophisticated assessments such as speech therapy or videofluoroscopy are unobtainable. Also, the conditions causing acute global cerebral depression that we described are not totally analogous to the more focal or chronic conditions referred to by Stanners, and some of the above reflexes may be attenuated in elderly or debilitated patients.

Though agreeing that a normal gag reflex does not necessarily indicate a safe airway, we think that the surprising feature of our data was the absence of the gag reflex in a group of fully conscious patients who would not normally be considered to be at any appreciable risk. We thus stand by our original assertion that in accident and emergency departments the gag reflex should be assessed independently of conscious level and used as an indicator of an airway at risk.

CHRISTOPHER MOULTON
A G PENNYCOOK

Accident and Emergency Department,
Glasgow Royal Infirmary,
Glasgow G4 0SF

1 Stanners AJ. Glasgow coma scale and gag reflex. *BMJ* 1991;**303**:1401. (30 November.)
2 Moulton C, Pennycook A, Makower R. Relation between Glasgow coma scale and the gag reflex. *BMJ* 1991;**303**:1240-1. (16 November.)

SIR,—A J Stanners would rely on videofluoroscopy in diagnosing aspiration in neurologically impaired patients.[1] This cumbersome investigation has largely been superseded in our hospital by direct observation of the larynx with a flexible fibreoptic nasendoscope while a bolus of milk is swallowed. "Milk nasendoscopy" is a quick and easy investigation that can be performed at the patient's bedside if required, with equipment and expertise that should be readily available in an ear, nose, and throat department. The patient is not irradiated, and the test can be repeated as often as required.

T J HOARE

Queen Elizabeth Hospital,
Birmingham B15 2TH

1 Stanners AJ. Glasgow coma scale and gag reflex. *BMJ* 1991;**303**:1401. (30 November.)

Chorionic villus sampling

SIR,—As chairman of the Medical Research Council's working party on the evaluation of chorionic villus sampling and coordinator of its European trial of chorionic villus sampling and amniocentesis[1] we are writing to dissociate ourselves from D T Y Liu's assertion, made in his capacity as a member of the working party, that "participants were still on the learning curve and results will differ if the trial is repeated."[2]

Our report explicitly commented on the absence of any evidence that the somewhat greater rate of fetal loss attributable to chorionic villus sampling declined as the trial proceeded, as would have occurred if those performing chorionic villus sampling were still on the learning curve. Only centres satisfying specific criteria regarding their previous experience took part in the trial. Furthermore, most patients were recruited in centres with a particularly high level of interest and skill in chorionic villus sampling. Whether results would differ if the trial is repeated is thus entirely speculative. Findings from several individual centres included the rate Liu expects based on his personal experience, and there was no evidence of heterogeneity between centres. Had Liu entered patients into the trial he would more convincingly have been able to compare his results with those of others.

It is also misleading to suggest that the Medical Research Council's European trial "was not sub-