The Journal of Bone and Joint Surgery

EDITORIALS

Controlled trials and meta-analyses can help resolve disagreements among orthopaedic surgeons

About a year ago, one of us sustained a fracture of the lower fibula while on holiday in New Hampshire. An orthopaedic surgeon there put the leg in a splint to support it during the transatlantic journey the following day, and explained that after the soft-tissue swelling had subsided, a plaster cast would have to be applied and worn for six weeks. A couple of days later, a British orthopaedic surgeon summarily dismissed his American colleague's prescription, arranged for a supportive bandage to be applied, and instructed his patient to "walk on the ankle until it hurts, and then a little bit more"!

Which of these orthopaedic surgeons was right? Were they both right, but interested in different outcomes of treatment? What were the consequences, in terms of short-term and long-term pain and function (and the costs of treatment), of acquiescing in the advice of the second rather than the first? And what is known about the effects of the various forms of physiotherapy which were prescribed subsequently? There was no response to a request for references to good review articles on the natural history and management of fractured fibula, and therefore, even if there are scientifically defensible answers to these questions, we remain unaware of them.

Had one of us had tennis elbow rather than a

fractured fibula, we would have been more fortunate because at least there is now an easily accessible article reviewing the various treatments offered to patients with this condition. On page 646 of this issue of the Journal, Labelle and his colleagues describe their systematic search for reliable evidence about the relative merits of the various ways of managing patients with lateral epicondylitis. They found that only a few of the reported studies came near to respecting widely accepted scientific principles, and that there is at present insufficient evidence for any of the current methods of treatment.

As Labelle et al point out, reliable evidence about the effects of this and many other aspects of orthopaedic care may only emerge through conducting properly designed randomised controlled trials. This requires an admission of uncertainty about the effects of the treatment or treatments in question (Collins et al 1992).

There are methodological challenges facing those who design trials of orthopaedic treatments, a matter which has been discussed in the Journal previously (Rudicel and Esdaile 1985). In particular, it is undoubtedly true that in the treatment of certain orthopaedic conditions, differences in surgical experience and skill must be important determinants of outcome. But, of course, this does not mean that these are the only important determinants of success.

For many common orthopaedic conditions (such as ankle fractures, tennis elbow, and hip arthritis) there is a rich variety of ideas about their management. The organisation by orthopaedic surgeons of properly controlled trials to compare those alternative forms of care currently being offered is likely to yield evidence which could be generalised to the benefit of many thousands of people every year.

Such trials should be designed to minimise the likelihood of bias – by ensuring randomised assignment to the comparison groups, by blinding the assessment of

I. G. Chalmers, MSc, FRCOG, Director, National Perinatal Epidemiology Unit
Radcliffe Infirmary, Oxford OX2 6HE, England.

R. E. Collins, MA, MSc, MB BS, Co-director
Clinical Trial Service Unit, University of Oxford, England.

K. Dickersin, PhD, Assistant Professor
Clinical Trials Unit, Department of Ophthalmology, University of Maryland, Baltimore, USA.

© 1992 British Editorial Society of Bone and Joint Surgery
0301-620X/92/5456 \$2.00
J Bone Joint Surg [Br] 1992; 74-B:641-3.

Data on pulmonary embolus and on mortality from all evenly randomized trials of perioperative subcutaneous heparin in orthopaedic, urologic or general surgery: 6366 patients allocated heparin (H) and 6426 allocated control (C)

Fig. 1

Reprinted, by permission of *The New England Journal of Medicine*, (318; 1162-73, 1988).

outcomes when this is feasible, by ensuring that all patients entered into a trial are accounted for in the analysis, and by reporting the results of all prespecified hypotheses. Bias must also be reduced by avoiding selective emphasis on a particular subgroup in a trial, or on a particular trial when several trials have addressed the same question.

There is another important source of confusion in assessing the effects of treatment, and that is random error. The dangers of being misled by random error can be reduced by specifying hypotheses in advance, before analysing the data, by minimising the number of interim analyses, by recognising that unpredicted 'statistically significant' differences may be due to chance (type I errors), and by reporting confidence intervals in addition to standard tests of statistical significance. Most importantly, study samples need to be large enough to provide adequate protection against the dangers of failing to pick up real differences when they exist (type II errors).

Ways of controlling biases and random errors need to be considered not only by those who design and analyse the results of primary research, but also by those reviewing evidence derived from a number of related studies. If several primary studies are judged to be sufficiently similar to each other, both biases and random errors can be reduced by considering their results in the context of a statistical overview, or 'meta-analysis'. Labelle et al had initially intended to use meta-analysis to reduce the risk of type II errors in the few controlled trials of tennis elbow treatments which were available. They found, however, that there were too many variations among the studies to allow the confident use of these methods.

Another attempt at meta-analysis of the results of

orthopaedic trials has recently been reported by Turner et al (1992), who studied 74 carefully selected papers on the outcome of surgery for spinal stenosis. They, too, came to the conclusion that "major deficits in study design and reporting were common, and these precluded firm conclusions". They end their report by urging investigators and editors to adopt more rigorous standards, and tabulate ten recommendations for the conduct of future studies.

Controlled trials of peri-operative subcutaneous heparin in patients undergoing surgery, by contrast, have not only been relatively well-designed (partly because they were easier to design well), but they have also been sufficiently similar for meta-analysis to be informative (Collins et al 1988). As shown in Figure 1, a meta-analysis of these trials reveals that heparin reduces the risk of postoperative death, an effect that none of the individual trials contributing to the meta-analysis had been large enough to detect reliably. These findings are of great relevance. The report of the confidential enquiry into peri-operative deaths in the UK in 1990 (Campling et al 1992) found that only 11% of their sample of 420 orthopaedic patients who died peri-operatively had received prophylaxis against thrombo-embolism, and concluded that this low prophylaxis rate was likely to have had a significant impact on the postoperative death rate.

In summary, orthopaedic surgeons can confidently expect to reap a harvest of evidence of great relevance to their patients: (i) by conducting reviews which respect scientific principles (using meta-analysis when appropriate) (Chalmers 1991); (ii) by recognising genuine areas of therapeutic uncertainty (when systematic reviews do not yield clear evidence and where substantial variations in practice exist between and within countries); and (iii) by collaborating in well-designed randomised controlled trials of sufficient size to address important questions when uncertainty remains.

Reliable meta-analyses will obviously be more easily undertaken (and updated) once a comprehensive register of orthopaedic trials exists (Dickersin 1992). There are encouraging developments along these lines. In May 1992, the Committee on Outcomes Studies of the American Academy of Orthopaedic Surgeons announced that it had established a Musculoskeletal Research Register. This register concentrates on studies which have assessed or are assessing outcomes of importance to patients, such as functional mobility and quality of life. For the reasons discussed above, it is focusing on randomised trials – published, unpublished, ongoing or planned. More information about the register may be obtained from Dr Robert Keller, American Academy of Orthopaedic Surgeons, 222 South Prospect Avenue, Park Ridge, Illinois 60068, USA.

More fundamentally, however, there is a need for wider acknowledgement that disagreements among orthopaedic surgeons should be resolved by well-designed

research. Neither the reputations of orthopaedic surgeons nor the interests of their patients are served by acquiescence in strongly held opinions which are unsupported by scientific evidence.

I. G. CHALMERS
R. E. COLLINS
K. DICKERSIN

REFERENCES

- Campbell EA, Devlin HB, Hoile RW, Lunn JN.** *The report of the national confidential enquiry into perioperative deaths 1990.* London: NCE-POD, 1992.
- Chalmers I.** Improving the quality and dissemination of reviews of clinical research. In: Lock SP, ed. *The future of medical journals.* London: British Medical Journal, 1991:127-46.
- Collins R, Doll R, Peto R.** Ethics of clinical trials. In: Williams CJ, ed. *Introducing new treatments for cancer: practical, ethical and legal problems.* Chichester: John Wiley, 1992:49-65.
- Collins R, Scrimgeour A, Yusuf S, Peto R.** Reduction in fatal pulmonary embolism and venous thrombosis by perioperative administration of subcutaneous heparin. *New Engl J Med* 1988; 318:1162-73.
- Dickersin K.** Why register clinical trials? - revisited. *Controlled Clin Trials* 1992:in press.
- Rudicel S, Esdaile J.** The randomized clinical trial in orthopaedics: obligation or option? *J Bone Joint Surg [Am]* 1985; 67-A:1284-93.
- Turner JA, Ersek M, Herron L, Deyo R.** Surgery for lumbar spinal stenosis: attempted meta-analysis of the literature. *Spine* 1992; 17:1-8.

Screening for congenital dislocation of the hip

It is now many years since the work of Ortolani (1937), von Rosen (1956, 1962) and Barlow (1962) ushered in the era of screening for congenital dislocation of the hip (CDH). It seemed at first that the problem of the late diagnosed established dislocation had been solved but, as more and more centres published their results, doubts began to emerge. Although the number varied, all reported some cases which had been missed by the screening programme.

Why were the early expectations not fulfilled? In retrospect it seems obvious that some of the criteria that we would now demand of a screening programme had not been met. These would include an accurate and repeatable test, an understanding of the natural history of the condition, and a safe and dependable method of treatment. The programme should also make economic sense. How does screening for CDH measure up? Any conclusions must be population-based, and it is significant that the poorest results have been reported by those who have studied a population rather than the captive occupants of a maternity hospital (McKenzie 1972; Williamson 1972; Jones 1977; Bjerkreim and Arseth 1978; Catford, Bennet and Wilkinson 1982).

Before the introduction of screening, the presumed incidence of established dislocation was 0.65 per 1000. It has probably risen since (Klingberg et al 1976; Catford et al 1982) and a generally accepted figure now would be around 1.5 per 1000. Yet, by clinical testing, 10 to 20 neonates per 1000 are considered to have abnormal hips. Clearly, we are overtreating this disorder.

When screening was introduced the natural history of the condition was ill understood. It was, for example,

assumed that dislocation arose by progression from an unstable hip but this is no longer certain (Berman and Klenerman 1986). Ultrasound studies have allowed the structure of the neonatal joint to be assessed better; it is now appreciated that a hip with normal anatomy can be unstable, and that one with very abnormal morphology can, nevertheless, prove to be stable (Clarke 1986; Tonniss, Storch and Ulbrich 1990). This recognition of our ignorance of the significance of the hip dysplasias must lead to a reappraisal of our expectations of clinical screening.

There is no point in screening unless a safe and effective treatment is available, one that will alter the natural history of the disease for the better. Overtreatment would not matter if it were risk-free but unfortunately it is not, and some cases of avascular necrosis have occurred in every series (Elsworth and Walker 1986; Bradley, Weatherill and Benson 1987).

Ultrasonic examination has the advantage over clinical testing in that it allows objective documentation of the anatomy of the developing hip in the neonatal period, and it is safe and non-invasive. It is particularly useful in selecting out those unstable hips that require treatment from those that do not (Gardiner and Dunn 1990) and in assessing their response to splintage. Its place in primary screening, however, has yet to be established. Skilled operators are required and the method has a low specificity in the first week of life, so that its use has been associated with high follow-up and intervention rates. It can be used selectively, for children considered to be at high risk, but children without risk factors may still present late (Clarke, Clegg and Al-Chalabi 1989).

Radiographs are of little value in the first few weeks of life and only begin to provide useful information at about three months of age, by which time adaptive changes have occurred. Although radiography may be useful as a safety net to diagnose those cases that have

G. C. Bennet, FRCS Ed, Consultant Orthopaedic Surgeon
Royal Hospital for Sick Children, Yorkhill, Glasgow G3 8SJ, Scotland.

© 1992 British Editorial Society of Bone and Joint Surgery
0301-620X/92/5473 \$2.00
J Bone Joint Surg [Br] 1992; 74-B:643-4.