Chapter 1

The nature of surgical evidence and evidence-based practice

John A Murie MA BSc MD FRCS, Consultant Vascular Surgeon and Honorary Senior Lecturer
The Royal Infirmary of Edinburgh, Edinburgh, UK

Introduction

Evidence is grounds for belief in something. Medicine in general, and surgery in particular, have not always followed an evidential path, but in recent times - the scientific era - evidence has been one of the great driving forces. Over the last few years it has become obvious that not all evidence is of equal value, and an evidential hierarchy has been suggested as a useful tool when considering the strengths and weaknesses of propositions. The following levels are the most well known:

- Ia - evidence obtained from systematic review of meta-analysis of randomised controlled trials;
- Ib - evidence obtained from at least one randomised controlled trial;
- IIa - evidence obtained from at least one well-designed controlled study without randomisation;
- IIb - evidence obtained from at least one other type of well-designed quasi-experimental study;
- III - evidence obtained from well-designed non-experimental descriptive studies, such as comparative studies, correlation studies and case studies;
- IV - evidence obtained from expert committee reports or opinions and/or clinical experience of respected authorities.

From the above hierarchy, recommendations for clinical practice can be grouped into the following grades:

- A - at least one randomised controlled trial as part of a body of literature of overall good quality and consistency addressing the specific recommendation (evidence levels Ia and Ib);
- B - well-conducted clinical studies but no randomised clinical trials on the topic of recommendation (evidence levels IIa, IIb, III);
- C - expert committee reports or opinions and/or clinical experience of respected authorities. This grading indicates that directly applicable clinical studies of good quality are absent (evidence level IV).

The above schemes are used throughout the chapters in this book and they are a real advance in evaluating the results of clinical science.

Another advance in assessing scientific work that has achieved prominence in recent years has been the introduction of the formal search strategy into the surgical literature. Not so long ago it was commonplace to come across a review article that contained a commentary based solely on those original papers that favoured the personal bias of the review author. Today, reputable journals will publish review work only if an assurance is given that all the appropriate literature has been competently searched and a clear explanation given of what has been included, what has been excluded, and the reasons why. This has isolated the old art form of the predudiced review to the realms of the medicolegal report.

The problem with researchers and authors

In an ideal world the literature would contain a full and proper description of all the studies ever carried out, but the scientific world is far from ideal. It is surprising how most readers accept the written word as reflecting ultimate truth, when a moment's reflection must surely convince even the most credulous that this cannot be so. Researchers and authors are human, and just as prone to insanity, perverse incentive and ignorance as everyone else. Gross fabrication of data might be uncommon, but it happens even at the highest level and outright lies appear in even the most prestigious journals. To support such a claim one need look only at the Woo Suk Hwang affair earlier this year [1]. This South Korean professor had achieved worldwide fame and near pop star adulation in stem cell research, choosing to publish his work in journals such as the *Lancet* and *Science*, but it all came to a sad end on the front pages of newspapers internationally when his claims to have used cloning techniques to create stem cell lines of 11 people were proved fraudulent.

It is probable that data mismanagement is far commoner than outright fabrication; in one study of scientific behaviour only 0.3% of researchers admitted to falsifying data while no less than 15.3% admitted to excluding observations or data points from analyses based on a gut feeling that they were inaccurate [2]. To this must be added failing to present data that contradict one's own previous research (6.0%), overlooking others' use of flawed data or questionable interpretation of data (12.5%), changing the design, methodology or results of a study in response to pressure from a funding source (15.5%), republishing the same data (4.7%) and so on. Very occasionally some of this behaviour can be attributed to the ignorance of a young researcher on how to conduct a scientific project, but far more often it is a matter of dishonesty for personal gain, be it money or prestige.

In addition to misconduct in data handling there is the problem of vested interest. Surgical instrumentation, implantable devices and drugs now play a greater part than ever before in healthcare, and their costs are high. With these increased costs has come an ever greater industry presence in medicine. Sometimes industry is an active partner in a research project, with full access to data and its interpretation, and with an axe to grind. More often industry is the silent partner, contributing in a general sense to departmental 'research' funds, and offering international travel and high quality accommodation to facilitate doctors' attendance at professional conferences. Is this likely to affect a surgeon's attitude to an instrument or prosthesis? Industry also specifically targets so-called 'opinion formers', easing their way around the world in business class seats to spread the word on how good the product is. Vested interest occurs in other forms too, but industrial connections, with their need for a return on an ever-enlarging investment, are the greatest cause for concern.

The problem with journals and editors

Surgical evidence appears mostly in journals. Journals, to remain healthy, have two concerns. First, they must attract papers; second, they must attract a readership. Why do authors choose to send work to one journal rather than another in the same field? The single biggest influence is perceived 'high quality', after which comes high impact factor. The impact factor of a journal is simply a measure of how often that journal's articles are cited. For a given (index) year, it is the number of citations made to articles appearing in the journal during the index year, divided by the number of articles published in the journal in the two previous years. Most editors feel, in general, that articles expressing a 'positive' result are more likely to appeal to readers (who are potential authors) and are more frequently cited than 'negative' articles (articles coming to no firm conclusion or expressing doubt about some approach). Such an attitude, of course, is recognised by researchers who, over time, are

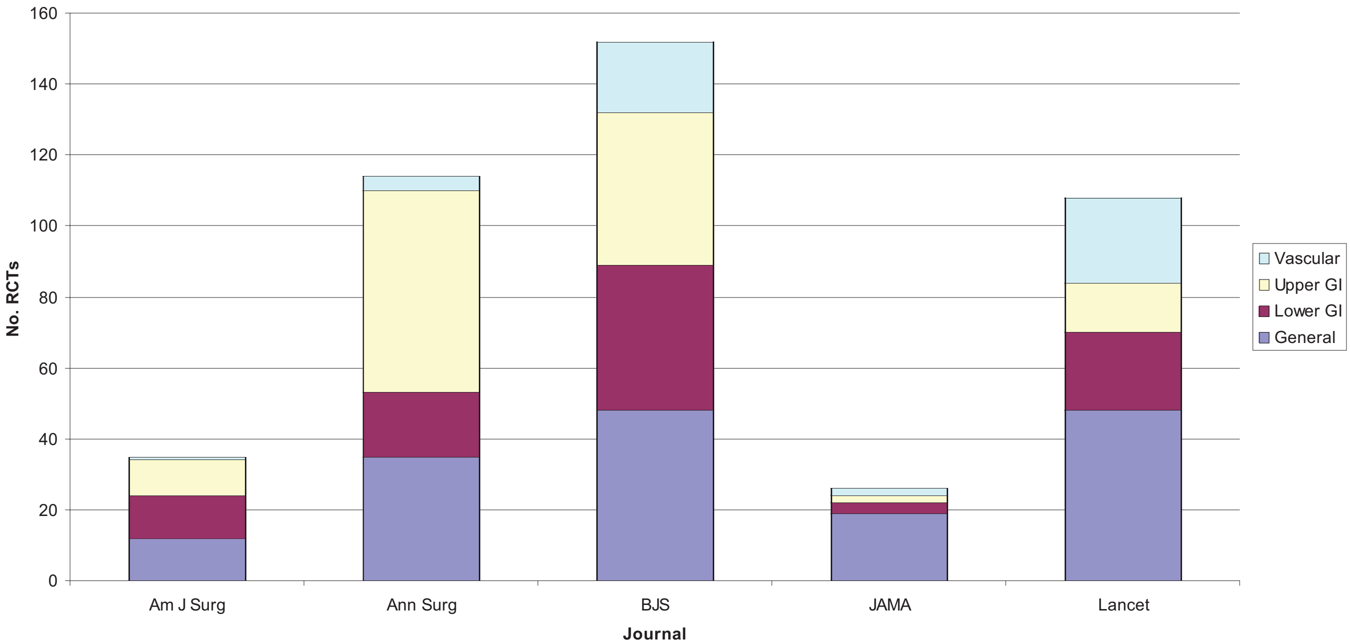

Figure 1. The number and type of randomised clinical trials published in major surgical journals 1998-2005.

effectively discouraged from submitting 'negative' work for publication, resulting in a positive feedback or vicious cycle. The result of this dysfunctional situation is that the literature is loaded with positive bias - inadvisable operations appear to produce wonderful results, prostheses perform with a brilliance that is stunning and drugs have perfect actions with few or no side effects.

Over the last couple of years the importance of the 'missing literature' has been recognised by editors and a plan is now afoot to redress the balance. Many of the most reputable journals now require all comparative prospective studies to be registered on a freely available website before they commence, if they are to be considered for publication at some future point [3]. This should ensure that there is a detailed record, free to all, of all such prospective work, including trials that produce a 'negative' result. The plan is not perfect, of course, as it is still likely to leave a positive spin within the pages of the journals, but at least reviewers will have access to unpublished 'negative' results and those conducting meta-analyses will be able to revert to authors for complete 'negative' or other unpublished data.

Evidence-based practice

Evidence-based medicine is usually defined as 'the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients' [4]. But is the evidence available adequate to make this notion a practical proposition? In broad terms only about 24% of surgical activity is based on level Ia/Ib evidence [5] and only 3.4% of all publications in the leading surgical journals are randomised controlled trials [6] (Figure 1). This represents a base of work that seems adequate to provide some guidance in the clinical sphere, but it is not a big base and, regretfully, many published guidelines reflect little more than level III/IV evidence. Clearly, there is room for improvement.

Even when high quality evidence is available, it rarely penetrates into clinical activity with ease. Many surgeons are resistant to change, others besotted with the latest gadgetry; personal opinion and hubris go largely unchallenged. Even in enlightened departments, dissemination of evidence-based guidelines has often done little to improve clinical practice, and some have concluded that the dream of expecting busy clinicians to follow this process is impractical [7, 8]. This has led to the synthesis of evidence into systematic reviews, development of computer-based clinical decision support systems, and improving access to useful information via the internet [9].

Such attempts to improve implementation of evidence-based practice are worthy, but they are unlikely to succeed without a change of departmental culture. Any success of evidence-based surgery will depend on the support of staff members, education of surgical trainees and regular practice within a patient-centred approach. Some hospitals hold an annual course for new staff and have weekly time set aside for discussion of evidence-based activity, using a standard format. 'Number needed to treat' and 'number needed to harm' are used rather than p-values, as the former are more informative measures of effect size. It is advisable to have one or two senior members of staff trained in evidence-based techniques who can take responsibility for the programme [10].

Conclusions

Surgical evidence and its effect on clinical practice depend on many factors. Some of these are sound, positive and encouraging, others are a cause of concern. The level of evidence hierarchy and the grading of recommendations are useful tools that focus the mind on the power of the available data. The recognition of evidence-based practice as a real entity is also important, as are attempts at its formal introduction into the clinical setting. Encouraging, too, is the fact that the 'missing evidence', which stems from failure to publish negative or equivocal results, has been recognised and is currently being addressed by the major journals. It would be laudable if the knowledge of how poor the surgical database actually is stimulates good research in the future [11].

References

1. Murphy TJ. On being downstream from faked scientific reports. *Br Med J* 2006; 332: 674.
2. Martinson BC, Anderson MS, de Vries R. Scientists behaving badly. *Nature* 2005; 435: 737-8.
3. De Angelis CD, Drazen JM, Frizelle FA, *et al.* Is this clinical trial fully registered? - A statement from the International Committee of Medical Journal Editors. *N Engl J Med* 2005; 352: 2436-8.
4. Sackett DL, Rosenberg WMC, Muir Gray JA, *et al.* Evidence-based medicine: what is and what isn't. *Br Med J* 1996; 312: 71-2.
5. Howes N, Chagla L, Thorpe M, McCulloch P. Surgical practice is evidence-based. *Br J Surg* 1997; 84: 1220-3.
6. Wente MN, Seiler CM, Uhl W, Buchler M. Perspectives of evidence-based surgery. *Dig Surg* 2003; 20: 263-9.
7. Guyatt GH, Meade MO, Jaeschke RZ, *et al.* Practitioners of evidence-based care: not all clinicians need to appraise evidence from scratch but all need some skills. *Br Med J* 2000; 320: 954-5.
8. McColl A, Smith H, White P, Field J. General practitioners' perceptions of the route of evidence-based medicine: a questionnaire survey. *Br Med J* 1998; 316: 361-5.
9. Applying evidence to health care delivery. In: *Crossing the Quality Chasm.* National Institute of Medicine. Washington, DC: National Academy Press, 2001; 145-63.
10. Ubbink DT, Legemate DA. Evidence-based surgery. *Br J Surg* 2004; 91: 1091-2.
11. Lara PN Jr, Higdon R, Lim N, *et al.* Prospective evaluation of cancer clinical trial accrual patterns: identifying potential barriers to enrolment. *J Clin Oncol* 2001; 19: 1728-33.