

The practical and ethical defects of surgical randomised prospective trials

Arnold Byer *Hackensack Hospital, Hackensack, New Jersey*

Author's abstract

This paper presents a strong criticism of the current enthusiasm for clinical randomised prospective studies in surgery. In the process, the author probes the 'intellectualism' or lack thereof in present day surgical attitudes. The subjects are examined against a framework of ethics and inescapable dilemmas. Ways of correcting the more obvious weaknesses are suggested.

The manuscript is, and is meant to be, provocative and is particularly aimed at the academic audience served by this journal.

In recent years there has been an increasing interest in clinical randomised prospective studies in surgery (CRPSS). The constantly recurring theme one hears is that if only this procedure, or method, or technique had been done in a randomised prospective fashion, the derived information would be more credible, if not downright irrefutable. The corollary presumably is that series not randomised and prospective can have little scientific value. Proponents of CRPSS leave, hanging over our heads like the sword of Damocles, the intimation that this will be the methodology of the future, the true measure of a series. The theory of CRPSS is that if a significant number of patients are treated with one method versus another, or not at all, useful information about that method will be forthcoming which will be definitive and of practical, as well as academic value. Undoubtedly this attractive, although not always benignly simplistic, approach to evaluating surgical management deserves some thought. And since it seems CRPSS is here to stay, I think it wise to reflect on what we may actually be doing, as well as on what we are being urged to give up.

It is hard to argue with a desire to arrive at valid and reproducible truths. I find it sad that that desire should be fuelled by dissatisfaction with more traditional methods. In fact, I suspect that it is more distrust than dissatisfaction that is at the heart of the matter. Unlike John Osborne, we can look back with considerable relief rather than anger, at the end of a century jam-

packed with scientific information and technological advances. It should be comforting to know that at least we are protected from committing the grosser follies. Our frustration with incomplete answers is not likely to lead us down the path of charlatanism. Having real reason to hope that we can be trusted to avoid the blunders committed in less scientifically fortunate times, why all the distrust?

Competition, ambition and dedication I prefer to consider as the most obvious, and amongst the more innocuous, traits of the human condition. Taken in context they are at least capable of stimulation and encouragement, the trademarks, if you will, of an exuberant, healthy curiosity. What gives me considerable concern is that this drive towards CRPSS may be the fifth horseman of an apocalyptic surgical fundamentalism which has no place in enlightened science. This fundamentalism or parallel movement to CRPSS, whichever you prefer, is characterised by the zeal to establish repositories of truth, a rigidity that is an unnecessary throwback to the era of The Great Surgical Schools and a poorly disguised arrogance, quick to brand as deplorable any and all so presumptuous as to differ. And, of course, one must document and document. No matter that this effusive documentation of some surgical subjects has arrived at a point where it is actually possible to sustain several totally opposing and conflicting points of view and document (from the literature) every single one. (Fortunately there are those who faced with this absurd situation are urging a re-evaluation of the whole subject). These attitudes, if not strictly meant to, do discourage, confuse, and, when backed by sufficient authority, stifle.

The impact of the scientific era as we know it bred empiricism as the basic scientific philosophy. Whether we like it or not, empiricism rests heavily on *a posteriori* reasoning in which the value of ideas and data are derived from experience; by a perceiving organism in a perceived universe as Hume put it. CRPSS could be a harking back to *a priori* reasoning. One of these two approaches is superior, but CRPSS assumes that it will tell us which.

Fundamental to scientific inquiry, usually, is the following scenario:

Key words

Clinical randomised prospective studies in surgery; clinical trials; medical experimentation; medical ethics.

- 1) Imagination, chance observation, reflective and logical thought give rise to an idea.
- 2) The idea is elaborated into a hypothesis.
- 3) The hypothesis is tested by collection of data (retrospective) that follow predetermined lines of investigation. The data are themselves collected from work with non-human models, animate and inanimate.
- 4) The acquired information is applied to actual problems.
- 5) Hard conclusions are derived from success or failure or any stage in between.

The *modus operandi* might be further subdivided into two major classifications: (a) Pharmaco-chemotherapeutic problems. Such studies have the advantage of back-up by chemical and physical analysis and proceed rather rapidly to animal research and double-blind clinical trials involving informed patient consent. Furthermore these studies have, in most instances, short end points, more or less rapid appearance of side-effects, ease of repetition, and a lot of available case material. More recent refinements, such as blood levels, immunological assays, etc, make frequent monitoring quite easy.

(b) Technological problems which are more pertinent to surgical studies. These involve instrumentation, synthetic and biological materials, and rather gross physiological and pathological concepts. The techniques themselves differ from individual to individual, as well as being subject to change by the same individual. Of major importance is case selection which involves personal preference, differing criteria, and highly individual judgment. In randomising the selection process, we must deal with unavoidable complications with difficult-to-evaluate effects on the primary goal of the study. These studies are further subject to limited case material, distant end points, and elimination of subjects because they are at a stage of disease when non-surgical treatment is clearly superior. The results are not easily or constantly reproducible, and the question of similarity is at best vague. Patient consent is either non-existent or at best rather loose. Traditionally, results have been evaluated by retrospective means, or in other words, 'Let's look at what used to happen, what should or could have happened, and what did happen.'

It would appear that the attraction of CRPSS lies in its attempts to emulate the ordered simplicity of the scientific inquiry of the first sort rather than in its ability to come to grips with the problems of the second. But face up to the latter we must.

Aside from the inherent problems of technological studies, CRPSS poses problems *per se*, some of which are contained within the very definition, and are so vague as to be abrasive to the scientific method: 'Randomised' – lacking aim or method, purposeless: 'Prospective' – likely to occur. As generally accepted and understood CRPSS follow a 'protocol' (list of contents?) In the kindest interpretation a protocol is a method of procedure and rules of study, or if one is

somewhat sceptical, a method suspect of orchestration.

The culmination of the study is a set of results and data collected either by the participants themselves, or more often by ancillary personnel who are told what to look for. The latter method is certainly less tedious, for the principals, and probably more objective, but would seem less complete in that chance observation, etc, would be eliminated.

With large studies, most people prefer that the results be refereed. Presumably the presence of referees has several purposes, the main one being to put into the hands of objective, recognised experts, the data collected, or more precisely the data the protocol calls for. The place of the referee is admittedly a hot spot. Rigorous diligence, on the one hand, robs the study of the flavour and astuteness of acknowledged experts, with all that their experience implies and, on the other hand, the referees' knowledge of and relationship with the participants is subject to accusations of less than optimum objectivity and outright politicking.

Surgical studies have such notoriously distant end points that new advances may make a particular study obsolete. Such a study may actually have a deterrent effect, being so demanding that the investigators are lulled into not seeking alternative treatment methods. If the protocol is rigid, then pertinent changes may not be incorporated. If flexibility is permitted, how much? And where flexibility permits changes to be incorporated is it the same study or has a deviant been produced, and how will a deviant affect the results? How valid are present-day statistical methods when considered against the magnitude of variables? Where does the role of teaching fit in? How are fellow collaborators selected and how does a constant changing of the guard affect the study?

The theory of CRPSS further presupposes that similar groups will be studied. If this were all it presupposed, the concept would immediately collapse. There is no such thing as 'similar groups' in the context of the desire to arrive at finite results. The usual criteria of age, sex, and a common disease process is an unpardonable gross over-simplification of similarity. Surgeons, of all people, should have no trouble understanding the necessity for fine-combing patient populations in order to come up with really similar groups and what a burden this is. We often imperfectly understand the disease process itself, nor do we at present have the facilities for evaluation of such factors as aetiology, genetic structure, concomitant pathology, socio-environmental factors, etc. Even if we did have these facilities close at hand, it is at best uncertain that we would always be dealing with the relevant problem at the same stage, evolution and probability of outcome.

I am unconvinced that randomising before the fact rather than after gives it any extra validity, although in the minds of some it may provide more respectability. Respectability, a much sought after social grace, is not necessarily a scientific virtue. It unhappily tends to achieve by rigidity and scorn what it cannot achieve by conviction and example. Evaluating experience can be

a painful undertaking. Yet, empiricism requires that we do so. In evaluating experience, it is at present fashionable to discard what does not fit the rigid format of accumulated numbers. Anything falling below that mark is dismissed as 'anecdotal' – a personal humorous experience – which in its current application has become a dirty word. Yet an anecdote is an experience, and the history of science is chock-full of anecdotes which provided some extremely good ideas. Newton did after all have a passing acquaintance with a falling apple. It was because of his genius, of course, that he was able to pursue this anecdotal observation and leave us a heavy legacy. (It is interesting that Newton's best work was completed in the years 1664–1666 at home when Cambridge University was closed.) Until we live in a more perfect world I favour the dissemination of some anecdotal information – it is, in any event, done all the time – having found that being considered anecdotal is an embarrassment that is short-lived and may in some way contribute to that branch of surgical curiosity usually referred to as 'Gathering Pearls'.

Case reports and small series are really larger anecdotes. It will not pass unobserved that their acceptance depends upon format and is in direct proportion to the *bona fides* and reputation of the observer-reporter. In short, it is difficult to understand these varying degrees of acceptance given the current deprecatory climate regarding anecdotes.

CRPSS and the rigidity of the protocol approach would appear to conflict heavily with the doctrine of *non nocere*. This doctrine carries the explicit and immutable obligation to provide the best treatment available within well-accepted norms of expertise. Any departures from this obligation, despite rationalisations about 'not really knowing what is best' has serious ethical consequences. (Who ever heard of a surgeon who really didn't know what was best?) The last surgical CRPSS I read presented the results of evaluating two operations for portal-hypertension. The authors reported a high mortality rate for the newer of the two procedures, but went on to explain that as they became more experienced their mortality rate declined. This was done in the face of large reported series (varied and multiple) which clearly outlined problems and pitfalls. It is not mentioned in the article to what degree the authors elected to prepare themselves for the performance of what was, for them, a new operation. It may be justifiable to accept a higher mortality and morbidity as a by-product of teaching residents. I, for one, have objections when this is the price to be paid for having CRPSS.

Informed patient consent in CRPSS, given the doctrine of *non nocere*, becomes more than just a ticklish matter. One would at least be duty-bound to demonstrate to prospective participants a tangible anticipated advantage, if not to the individual personally, then at least to society at large. Surgical studies already have the additional disadvantage of being concerned in unequal parts with side-effects, complications (common to that particular surgical sector) and cause and effect,

without adding literally insult to injury. To expose patients to unusual risks in the pursuit of a nebulous academic goal is ethically untenable.

The late C P Snow has discussed the problems inherent in the different languages spoken by the literary and scientific communities. Serious attention to his admonition that we develop a common language may well provide the key, or at least a clue, for resolving this delicate dilemma of the need to know in an informative context versus the need to study.

I see very little that is of real merit in CRPSS as we know them today. I suspect that we have not seen a proliferation of them because their shortcomings are as apparent to others as they are to me. Is there something we can do that would be better? I hope so and put forward the following brief suggestions:

- 1) The establishment of small study groups which would:
 - a) Establish the rules.
 - b) Select referees or themselves act as referees.
 - c) Promote and invite the broadest possible co-operative studies, farming out segments of the study to participants with interest and proven experience in a given area.
 - d) Authorise referees to carry out spot-checks to pick up trends. They would then determine whether to alter segments of the study or pursue new avenues.
- 2) The studies to include concomitant pertinent research projects that significantly affect the study.
- 3) Written patient consent and ensured follow-up.

Wherever this may lead, the addition of a healthy realisation that the outcome may not be the last word is important. 'To strive, to seek, to find, and not to yield', (to borrow Tennyson's words) honestly and industriously, occasionally muddling through, occasionally showing flashes of brilliance is not the worst of all possible worlds. It has at least brought us this far.

Commentary

Michael Baum, *Department of Surgery, King's College Hospital Medical School*

I wish to thank the editor for allowing me this opportunity to comment on the paper by Dr Arnold Byer entitled 'The practical and ethical defects of surgical randomised prospective trials'. Dr Byer's somewhat self-indulgent style of prose fails to disguise the fact that his paper is a mere polemic, using that word in its original meaning (a method of conducting theological controversy). For this reason I was at first very hesitant to respond, recognising the dangers in crossing tautological swords with one of the champions of non-

Key words

Randomised controlled trials; clinical randomised prospective studies in surgery; clinical trials; medical experimentation.

science. These debates tend to share the futility of argument with religious bigots, Marxists and psychoanalysts! Nevertheless I feel compelled to respond in a spirit of sorrow rather than anger at seeing the forces of intellectual reaction marshalled behind a smoke screen of literary pretensions. (Two can turn out the purple prose as well as one.) If nothing else I can at least draw readers' attention to the recent bibliography showing how the 'less scientifically fortunate' doctors of the Aristotelean/Galenic persuasion are still alive and well and found walking the marbelled corridors of modern high technology hospitals.

Rigidity and arrogance are certainly not attributes of the proponents of randomised controlled trials (RCT) who by their very adherence to scientific empiricism are prepared to say 'I don't know'. It is the advocates of non-science or conceptual rationalism who are guilty of rigidity and arrogance because having built their conceptual framework they know it to be perfect and are not prepared to subject their ideas to scientific tests of refutation. The traditional methods for arriving at the truth that Dr Byer apparently espouses are the anecdotal approach and the use of historical controls. I provide a list of recent publications in the reference section which illustrate the inherent danger of such an approach (1,2,3,4,5,6) and would like to remind readers that the use of anecdotal evidence alone is the province of the practitioners on the fringe. If medicine does not strive to be a science then medical practitioners cannot be critical of the quacks when they advocate Laetrile and coffee enemas as treatment for cancer.

This now leads us on to Dr Byer's doctrine of *non nocere*. Apart from the tragic and unnecessary mutilation resulting from radical surgery based on an uncritical acceptance of Halstedian dogma (1) let us remember some of the other harmful legacies of 'muddling through'.

For 1600 years pus in a healing wound was considered laudable and treatment was even designed to induce suppuration which was considered an integral component of the healing mechanism (4). Coming to more recent times, nephropexy was a fashionable treatment for back-ache for about 40 years until it passed out of fashion in the 1930s (5). Again this approach was founded on anecdotal evidence that appeared to corroborate the hypothesis that back-ache was due to drooping kidneys.

More recently still, let us remember the thousands of babies blinded by retrolental fibroplasia as a result of high oxygen concentration in their incubators (6). The initiators of this form of therapy for respiratory distress syndrome in neonates would have considered it quite unethical to subject the new treatment to the rigours of a RCT but we must never forget that nature is unpredictable and it is almost inevitable that any innovation in medical treatment will induce unexpected and unpredictable harmful side-effects. Currently, I am heavily involved in RCTs of adjuvant systemic therapy for cancer, a truly wonderful idea,

and many medical oncologists in the United States of America feel it unethical to do RCTs because they already know by supernatural precognition that the treatment works (because it ought to work). Yet applying the principle of *non nocere* we have to recognise the immediate toxic side-effects which have been accepted in anticipation of a net benefit in survival. Already the early promising trends are beginning to disappear, once again confirming the contrariness of nature (7).

Dr Byer also feels that it is acceptable to wait for the flashes of brilliance from the individual clinician that by quantum leaps will solve all medical problems. Yet sadly, the history of our profession has demonstrated again and again that the rigid arrogance of the conceptual rationalist blinds him to this flash of brilliance. Thus Lister's discovery of antiseptics was ridiculed in our country and took nearly 40 years to be accepted whilst at least one generation of surgeons were continuing their laudable efforts to induce laudable pus (4).

Finally, Dr Byer feels that there is some kind of dilemma facing a surgeon who does not know what is best for his patient. For once I agree with him yet I for one am proud to admit that I do not know best and cope with this dilemma by conveying an aura of confidence for the sake of my patients whilst being plagued with nightmares of uncertainty. In the United Kingdom I am not considered an eccentric, because the majority of my generation of surgical colleagues feel the same way. I know things are different in the United States of America but that is the fault of their system of surgical education. In America surgeons are technologists first and doctors second whilst in this country surgeons are physicians first who occasionally use surgical intervention in the overall management of those diseases that by tradition have been referred for the knife.

To paraphrase Bertrand Russell – to progress in the science of medicine we have to learn to live with doubt whilst avoiding the paralysis of inaction – that is what RCTs are all about.

References

- (1) Baum M. Scientific empiricism and clinical medicine: a discussion paper. *Journal of the Royal Society of Medicine* 1981; 74: 504–509.
- (2) Vere D. Controlled clinical trials: the current ethical debate. *Journal of the Royal Society of Medicine* 1981; 74, 2: 85–88.
- (3) Sacks H, Chalmers T C, Smith Jr H. Randomised versus historical controls for clinical trials. *The American journal of medicine* 1982; 72: 233–240.
- (4) Linder F, Forrest H. The propagation of Lister's ideas. *Surgery, gynaecology and obstetrics* 1968; 127: 1081–1086.
- (5) Barnes B A. Chapter 8: Discarded operations: surgical innovation by trial and error. In: Bunker J P *et al*, eds. *Costs, risks and benefits of surgery*. New York. Oxford University Press, 1977.
- (6) Silverman W A. *Retrolental fibroplasia: a modern parable*. London: Academic Press, New York: Grune & Stratton, 1980.
- (7) Berstock D A, Baum M. Breast cancer: adjuvant therapy. *Clinics in oncology* 1982; 1, 1: 177–190.