

EVIDENCE-BASED MEDICINE

Never Better than the Evidence Itself By Nick Lane, PhD



Nick Lane

Evidence-based medicine is all-important these days. If the British Journal of Ophthalmology is anything to go by, ophthalmologists can hold their heads pretty high: a study published last December showed that an impressive 77% of 274 consecutive interventions in the Hong Kong Eye Hospital were evidence-based – a figure that compares well with other fields of medicine. But there is a deeper malaise, not just in ophthalmology but also across medicine in general – the quality of the evidence base. Responsible clinicians may strive to follow it, but is it any good?

BACKGROUND NOISE

The first problem with finding good evidence is the level of background noise – enough to deafen all but the most intrepid seekers after the truth. I'm reminded of a passage in the Hitch Hikers Guide to the Galaxy, in which a population of gifted beings is cursed with the gift of telepathy. The only way they can prevent themselves from going mad is to chatter incessantly about trivia, at the top of their voices. It's hard at times not to look on the plethora of clinical data competing for our attention as a similar cacophony. It's all written in the same language of data and statistics, it all looks scientifically persuasive, at least to a cursory glance, and it's probably contradictory.

We need to look beneath the polished veneer of statistical evidence to make a judgement on the quality of information, and this is not always easy to do. Nor are ophthalmologists necessarily trained to do so. A study in the BMJ in 2000 showed that, in the UK, ophthalmology was one of five major specialties (out of 15) in which the Royal College exams did not assess skills in evidence-based medicine.

“We need to look beneath the polished veneer of statistical evidence to make a judgement on the quality of information”

Part of the problem is the sheer accessibility of information, which amounts to a kind of electronic telepathy. Look anything up on PubMed these days and it's surprising if you don't turn up 7000 hits. Refine your search terms, and who knows what you might be missing.

But try specifying 'randomised clinical trials' (RCT) and you may be surprised to find that your 7000 hits has just melted away to seven, of which the most recent was published back in 1997. The best of these trials probably randomised a total of 21 patients into three groups, and the investigators were unable to reach a firm conclusion. If there's a p number, you'll be lucky if it grazes 0.05. And frankly, with so few patients, if it's much better than that, you should be worried about the ethics or the bias of the study.

In the final paragraph of the discussion, the authors probably called for a larger study to test their conclusions. There's nothing wrong with that – but it's rare to see the follow-up trials actually done. After all, there's now an RCT in the literature, so why bother? It's no longer new, and research moves swiftly on. Quite a few of the RCTs cited in the reassuring BJO study are of this type: small,

equivocal trials, to put it politely. A large pile of them doesn't necessarily add up to better evidence.

It's reasonable to conclude that there is a serious problem in ophthalmology. There are a handful of randomised trials that stand comparison with any in medicine, but these underline the paucity of well-designed, well-performed trials in other areas. Consider, for example, the AREDS trial, the ETDORS trial, the TAP and VIP trials for PDT, the ongoing ESCRS endophthalmitis study and the ongoing memantine trial in glaucoma. All were (or are) large enough to detect a real difference, enrolling a few hundred, often a few thousand, patients. The patients were chosen according to careful selection criteria, taking into account the natural history of the diseases.

Apart from the inclusion criteria, the relevance of outcomes is critical – not just the measures themselves but also the timing. The trials mentioned lasted for several years, a respectable time to measure reliable outcomes in predominantly degenerative eye conditions. All were relevant to patients, with an array of complementary endpoints, from visual acuity changes, to IOP reductions, or angiographic analyses, to Quality of Life. Certainly they have some limitations, but in general their findings can be trusted: this is solid research.

How rare these examples are can be seen from a quick visit to the National Guideline Clearinghouse (www.guideline.gov) or the Cochrane Eyes and Vision Group (CEVG) (www.cochraneeyes.org).

Take the guidelines first. Most guidelines these days give an indication of the quality of evidence. The exact schemes vary according to the organisation compiling the guidelines, but a typical scheme describes well-designed RCTs as Level I, cohort or case-control studies as Level II, and case reports or expert opinions as Level III (see Box 1).

If you stray into the AAO guidelines on cataract in the adult eye, you'll be lucky to find a single piece of evidence that rises above Level III. The same applies to the recommendations of the AAO on refractive errors. Even the recommendations for primary open-angle glaucoma rarely merit a rating above III; the only recommendations that receive a Level I are an IOP check within 30 to 120 minutes of surgery, and the use of topical corticosteroids in the postoperative period (unless contraindicated).

The problem is made more explicit when you visit the Cochrane Library. Take a look at surgery for age-related cataract again. In their 2002 systematic review (the most recent), the Cochrane reviewers found just six trials that satisfied their standard criteria, randomising a total of 7,828 people. Of these, only a single trial, randomising 476 patients, compared phaco with extracapsular surgery. Although the trial showed that phaco gives a better visual outcome, the Cochrane reviewers cautioned against extrapolating to settings outside a specialised hospital environment in developed countries (the trial was performed at the Institute of Ophthalmology at UCL in London by Minassian et al.). Similarly, only one trial, the

RATINGS OF STRENGTH OF EVIDENCE

(from the American Academy of Ophthalmology)

I. Level I includes evidence obtained from at least one properly conducted, well-designed randomised controlled trial. It could include meta-analysis of randomised controlled trials.

II. Level II includes evidence obtained from the following:

- Well-designed controlled trials without randomisation
- Well-designed cohort or case-control analytic studies, preferably from more than one centre
- Multiple-time series with or without the intervention

III. Level III includes evidence obtained from one of the following:

- Descriptive studies
- Case reports
- Reports of expert committees/organisation
- Expert opinion (e.g., Preferred Practice Pattern Panel consensus)

Madurai Intraocular Lens Study (MIOLS), was powered to show that ECCE/PC-IOL provides better visual outcomes at one year than ICCE-AG. They recruited 3,400 patients.

The most striking aspect of the Cochrane reviews is the small number of trials considered to be valid in many areas of ophthalmology – often no more than a handful, despite the hundreds of presentations at scores of conferences every year. There is a constant refrain of phrases like "Only three studies focussed on patient relevant outcomes", or, from Richard Wormald, the Coordinating Editor of the CEVG (referring to glaucoma): "There are numerous uncontrolled case series discussing modifications of the procedure and different types of implant, but the quality of evidence remains poor".

FAIRGROUND NOISE

It's perhaps unkind to criticise the efforts of individuals to put together case series: ophthalmologists are doing their best to generate coherent evidence from their own practice, and subjecting their results to the rigour of peer review. Certainly this is far better than not keeping records, not following up, not considering patient outcomes, not sharing best practice, not undergoing peer review. But I claimed at the start of this piece that the problem reflects a deeper malaise in medical research: the operation of a market that has the potential to undermine objectivity. I have nothing against market forces in medical research, or commercial gain in science, but they can, and do, lead to some shady practices.

Researchers are under a huge pressure to publish. Careers are built on a good publication record. The journals compete among themselves for high impact ratings. They put out their own press releases to draw the media's attention to 'revolutionary' breakthroughs, often prematurely. Press releases can give a misleading impression of the real value of a paper, and are embargoed until publication date, enabling journalists to report the story before the scientists have had any chance to digest its import. Naturally, the press tends to be interested in the sexier end of medicine – new technologies that border on science fiction. As Donald Minckler

put it at the last AAO annual meeting: "A number of 'Roman candle' innovations have been introduced into ophthalmology with great fanfare, only to fade quickly from view. Many innovations are touted without sufficient evidence of their safety and effectiveness derived from RCTs."

“A number of 'Roman candle' innovations have been introduced into ophthalmology with great fanfare, only to fade quickly from view” - Donald Minckler

These pressures distort the objectivity of data. Journals tend not to publish negative findings, which already biases the evidence base. If you want to be published, you'd better find something positive to say. How? Well, the BMJ can help. In the last Christmas issue, the BMJ ran a series of spoof articles on how to conceal the truth in clinical research, including How to make a compelling submission to NICE: tips for sponsoring organisations; and HARLOT plc: an amalgamation of the world's two oldest professions. The acronym HARLOT said it all: How to Achieve Positive Results without actually Lying to Overcome the Truth. And some of the advice cut unpleasantly close to the bone – compare your intervention with inactive or toxic controls, or dose the competitor's drug inadequately; select only patients who fail to respond to standard treatment, and use that as the comparator; define the withdrawal criteria to enable the exclusion of patients who find the drug toxic

from the intention-to-treat analysis; combine outcome measures until one combination achieves high significance; use only relative risk reductions (not absolute risk reductions) and avoid citing actual patient numbers, to prevent re-analysis, and so on.

Deeply cynical, but I challenge anyone who spends their days considering the meaning of trial results to tell me that they haven't come across examples of this kind of thing themselves.

In this context, I find it perversely encouraging that so many large pharmaceutical-sponsored trials do end in failure. Take memantine, for instance, the neuroprotective agent currently being evaluated in a four-year trial in patients with glaucoma: it is the sole survivor of the numerous NMDA antagonists that underwent trials for stroke, Alzheimer's disease, Parkinson's disease, and so on, in the 1990s (and the only one to be licensed in Europe for Alzheimer's disease). All the other NMDA antagonists failed in late-stage clinical trials, usually because of their negative risk-benefit profile. It was bad news for the companies concerned, and the patients who might have benefited, but good news for the health of medical research.

Another area in which potential troubles abound is the ghost writing of clinical papers. I have nothing against ghostwriting in itself: why not employ a professional writer to get important results into the clinical domain as quickly and efficiently as possible? The trouble is that the use of ghostwriters is not acknowledged in the small print, let alone in the main author list.

I recall reading an anguished commentary from David Sharp, then the deputy editor of The Lancet (now retired), who had just attended the European Medical Writers Association annual conference. He expressed surprise at the existence of ghost writers, but anyone who has worked with pharma

companies knows that a major product is supported by 'publications plans', in which hundreds of clinical papers are ghost written, and targeted to key audiences in specific journals at strategic times. Again, there is nothing wrong with this if it helps disseminate important findings quickly to the people who need to know; but it would be interesting to conduct a straw poll among clinicians to find out how many know they are being targeted in this way. While virtually all of these papers are written professionally and contain valid data and interpretations that have been carefully reviewed by the 'authors', it is nonetheless a deception, and blurs the distinction between science and marketing.

So what should be done? In ophthalmology, too many case-control series are undertaken by small teams with little backing and ultimately little validity. What do they add if they are dismissed as valid evidence by balanced arbiters, such as the Cochrane reviewers? On the other hand, larger trials need to be carefully designed by unbiased researchers, probably from a number of institutions, and perhaps under the auspices of organisations like the ESCRS, (see Box 2). Ophthalmologists can be proud of a number of such trials, some of which I mentioned; but this is the level to which all should aspire. Evidence-based medicine can only ever be as good as the evidence it's based on.

APPLIED TO EVIDENCE-BASED MEDICINE

Applying Evidence-Based Medicine to your own Practice

- Check the Cochrane reviews and any available guidelines; stick to systematic reviews, meta-analyses and RCTs when evaluating evidence if possible – don't value an uncontrolled case series higher than an RCT, even if the results look impressive
- Always check on patient numbers – if there's less than 20 to 30 in a trial, the results can't be any better than 'promising' and shouldn't sustain any claim
- Always check that the patient recruitment makes sense, that the outcomes are relevant, and that you're not being misled by an unsubstantiated big claim
- Check the comment links on PubMed – you can learn a lot from other people's criticisms of trials: it helps to see where potential biases lie

Designing your own Trial

- Try to collaborate with other centres interested in the same problem, rather than just doing another case-control series by yourself
- Spend time planning relevant recruitment, relevant endpoints, and follow up your patients for a meaningful period
- Consider relevant comparators, and try to power your trial so it can determine a difference – make sure you enrol enough patients and work with a medical statistician if possible
- Register your research with the Cochrane Eye and Vision Group – and get their advice on your study

Publications – How to Write it Up

- Include relevant data in the abstract, including patient characteristics at baseline, duration of trial, and relevant outcomes; omit data that should be in the methods section, about laser settings, etc
- Always give patient numbers, standard deviations, intention-to-treat analyses, and absolute risk reductions – if you've got nothing to hide, don't hide it! Make it possible for systematic reviewers to incorporate your data into best practice guidelines
- State your main conclusion clearly in the abstract if you want anyone to follow up on your full paper, and make sure it follows logically from the data you presented in the abstract