CLINICAL PRACTICE

Evidence-based medicine – are we boiling the frog?

D J J Muckart

David Muckart is Associate Professor of Surgery at the Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa, and Chief Specialist at the Level I Trauma Unit and Trauma Intensive Care Unit, Inkosi Albert Luthuli Central Hospital, Durban. His clinical and research activities concentrate on the management of the critically injured, and he has a keen interest in the history of medicine and surgery.

Corresponding author: *D J J Muckart (davidmuc@ialch.co.za)*

Evidence-based medicine has been defined as 'The conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.' There are two major assumptions in this statement. First, it is assumed that the evidence is in fact the best. Unfortunately this is not necessarily so, and published evidence is affected by bias, sponsorship, and blind faith in mathematical probability which may not be clinically relevant. Second, the evidence is population based and may not be applicable to the individual, and blind adherence to this concept may cause harm. We must not abandon clinical experience and judgement in favour of a series of inanimate data points. Medicine is an uncertain science.

S Afr Med J 2013;103(7):447-488. DOI:10.7196/SAMJ.6805

For more than 2 000 years, anecdotes, personal experience and bias dictated medical practice. Untold harm was caused by unsubstantiated proclamations such as that by Dupuytren, who held that under no circumstances could a structure as insignificant as the appendix be responsible for any abdominal mischief.[1] The traditional hierarchical training structure, whereby the consultant's word was law, perpetuated such dogma.

Medical practitioners have now aligned themselves with their legal colleagues and a scientific standard of proof, based on best available evidence, is required to substantiate current practice. Unlike the legal system, however, scientific proof must reach a level of certainty greater than 'beyond reasonable doubt'.[2] This has resulted in the construction of the evidence pyramid, but as with legal argument, the evidence provided by each level has been contested, strong opinions being voiced by opposing camps.[3,4] Even when performed with appropriate numbers of patients, assignment and blinding, bias may confound the best of randomised controlled trials.

Publication bias

A publishing bias against studies with negative or inconclusive findings exists.[5-7] Clinical trials in which the results show a significant difference are three times more likely to be accepted, and are likely to be published more rapidly, than those with insignificant findings. The exclusion of unpublished data may skew the findings of any meta-analysis. Furthermore, results that are detrimental to the tested product may even be deliberately suppressed by the manufacturer.[8-12]

Sponsorship by for-profit organisations

An analysis of 159 trials involving 12 different specialties concluded that there was a significant finding in favour of the trial drug if the study was funded by for-profit organisations, which could not be explained by methodology, statistical analysis or type of study.[13] A similar review found that 51% of studies funded by for-profit organisations were in favour of the trial drug, compared with only 16% of studies sponsored by non-profit organisations.[14] This must cast doubt upon the validity of certain conclusions. As stated by Angell (editor of the *New England Journal of Medicine* for two decades), 'Physicians can no longer rely upon the medical literature for valid and reliable information.' She reluctantly concludes that prescription drugs are not nearly as effective as the publications on randomised trials suggest.[11] An analysis of highly cited trials published in the three journals with the highest impact factors (*New England Journal of Medicine*, *Lancet*, *Journal of the American Medical Association*) and those with an impact factor greater than seven, showed that 30% of trials initially reporting highly significant positive findings were found in subsequent studies to either overestimate treatment effect or show no benefit.[15,16] The effect of funding extends beyond drug or equipment trials. Guidelines and consensus statements by panels of experts are frequently supported by industry, and the members of such panels may have financial affiliations with the sponsoring company.[11,17]

Ghost and guest authors

Ghost authorship takes two forms. In its benign form, professional medical writers may improve a manuscript without altering its scientific content. They may be acknowledged in the text but will not appear on the list of authors. A more malignant tendency has spread in industry-sponsored studies: the initial draft is compiled by company employees, before academically affiliated authors, often regarded as key opinion leaders, are sourced as principal or second authors without having substantially contributed to the study.[18]

Data fabrication

From painted mice to post-op pain relief, instances of trial misconduct and data fabrication have raised their ugly heads. This is cause for serious concern and casts a shadow over medical evidence. A recent analysis found that 2% of scientists admit to fabricating or modifying data at least once, and one-third confess to questionable research practices. Interrogating colleagues revealed more alarming figures of 14% for data falsification and 72% for debatable scientific behaviour.[19]

Clinical versus statistical significance

The keystone in the bridge between clinical trials and conclusions is statistical significance. Simply put, it produces a mathematical probability of whether the results of a study comparing two or more groups are due to chance, a 5% risk of the results being falsely positive

deemed acceptable. As amusingly described by Hall,[20] this is not due to divine intervention but was the learned opinion of the statistician Fisher. In essence, it is therefore based on subjective expert opinion, the antithesis of evidence-based medicine.

Statistical significance, however, may have little to do with clinical relevance and must not be confused with biological importance.[20] Luus *et al.*[21] suggest that clinically relevant differences and statistical significance concur only by coincidence. They emphasise that although clinicians need not be conversant with statistical methodology they should understand the results, and statisticians must have some understanding of the clinical problem in order to generate statistical results commensurate with meaningful clinical conclusions. Trials aim to determine whether the aspect under scrutiny will affect clinical practice, so results should be expressed in clinical, not mathematical, terms. The latter are very susceptible to sample size, and meaningless clinical differences may be statistically significant. The reverse also holds true if the sample size is too small. In his book *The Last Well Person*, Hadler[22] argues that no study can control for all confounders, and an absolute difference of less than 2%, even if mathematically significant, should be viewed with caution.

Of greater clinical relevance is the number needed to treat (NNT), the reciprocal of absolute risk reduction, which defines how many patients need to be treated for one to gain benefit.[23] This calculation has no correlation with probability values, but gives an assessment of clinical impact. Of equal or greater importance is the number needed to harm (NNH), which assesses the possible adverse consequences of a particular intervention. The POISE (Post Operative Ischaemic Evaluation) study epitomises these concepts.[24] This is the largest randomised controlled trial to assess whether the risks of postoperative cardiovascular events can be lowered by peri-operative beta-blockade. A highly significant reduction in non-fatal myocardial infarctions was found in the treated group, with an NNT of 66. The incidence of stroke doubled, however, with the NNH being 200. For every three patients spared a cardiac event, one would potentially suffer a cerebral insult. Among those in the treatment group who suffered a stroke, only 15% regained full function, and 26% were left severely incapacitated. The choice between the risk and sequelae of a non-fatal myocardial infarct versus a disabling stroke is a matter of clinical judgement and patient preference, not mathematical probability.

Errors in clinical trials

Errors in clinical trials may be random or systematic. The former is unpredictable and may skew data both positively and negatively. An increase in sample size reduces its occurrence. Systematic error is not eliminated by increasing the sample size, and arises when a trend in the data occurs that is actually false. This results from three types of bias, namely selection, misclassification and confounding. Selection bias occurs when a test is inadvertently skewed to favour a subset of patients. Misclassification bias describes the error of placing patients in an incorrect category, resulting in a heterogeneous rather than homogeneous population under scrutiny. This is especially true where standard therapies are normally titrated against physiological end-points rather than fixed dose regimens. Deans *et al.*[25] cite the acute respiratory distress syndrome low tidal volume trial as a prime example; patients were randomised to fixed tidal volumes of either 6 ml/kg or 12 ml/kg, whereas the standard practice would be to titrate treatment in accordance with airway pressures and compliance. The identical scenario pertains to transfusion triggers. Younger patients without coronary artery disease may tolerate a lower haemoglobin level than the elderly cardiopath, and conversely, overtransfusion in the young may have a detrimental effect.[26] Such insufficient or excessive therapy may contribute substantially to differences in the trial results. Confounding bias refers to the mistaken relationship found between two variables because of a third unaccounted factor.

The boiling frog

There is a physiological anecdote that if a frog is placed in boiling water, it will leap out immediately. If the water is initially tepid, however, and slowly heated to boiling point, the frog will remain until boiled alive. This example has been used in various scenarios, including economics and global warming, to illustrate the concept that slow change may pass unrecognised until harm occurs. From initially tepid waters the zeal for evidence-based practice has now reached boiling point, and if the shortcomings are not appreciated, evidence-based medicine may itself become a boiled frog. The concept is disease and not patient orientated, is not scientifically perfect, and must not be viewed as exclusive.[27] As Osler observes, 'Variability is the law of life and no two individuals react alike and behave alike under the abnormal conditions which we know as disease. The good physician treats the disease, the great one treats the patient.'[28]

As in criminal law, even if the evidence is beyond reasonable doubt, it is rarely unequivocal or indisputable; evidence is not synonymous with truth. Even in modern practice the aphorism of Osler still holds true, 'Medicine is a science of uncertainty and an art of probability.'[29]

1. Cartwright FF. The Development of Modern Surgery. Liverpool: C Tinling & Co., 1967:204.
2. Miller DW, Miller CG. On evidence, medical and legal. Journal of American Physicians and Surgeons 2005;10(3):70-75.
3. Karanicolas PJ, Kunz R, Guyatt GH. Point: Evidence-based medicine has a sound scientific basis. Chest 2008;33(5):67-70. [http://dx.doi.org/10.1378/chest.0068]
4. Tobin MJ. Counterpoint: Evidence-based medicine lacks a sound scientific basis. Chest 2008;133(5):1701-1774. [http://dx.doi.org/10.1378/chest.0077]
5. Stern J, Simes RJ. Publication bias: Evidence of delayed publication in a cohort study of clinical research projects. BMJ 1997;315(7109):640-645. [http://dx.doi.org/10.1136/bmj.315.7109.640]
6. Montori VM, Smieja M, Guayatt GH. Publication bias: A brief review for clinicians. Mayo Clin Proc 2000;75(12):284-288. [http://dx.doi.org/10.4065/75.12.1284]
7. Gregor S, Maegele M Sauerland S, Krahn JF, Peinemann F, Lange S. Negative pressure wound therapy: A vacuum of evidence? Arch Surg 2008;143(2):189-196. [http://dx.doi.org/10.1001/archsurg.2007.54]
8. Psaty BM, Kronmal RA. Reporting mortality findings in trials of rofecoxib for Alzheimer disease or cognitive impairment. JAMA 2008;299(15):1813-1817. [http://dx.doi.org/10.1001/jama.299.15.1813]
9. Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B. Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: Review of studies in new drug applications. BMJ 2003;326(7400):1171-1173. [http://dx.doi.org/10.1136/bmj.326.74001171]
10. Turner EH, Matthews AM, Linardatos E, Tell R, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 2008;358(3):252-260. [http://dx.doi.org/10.1056/NEJMsa065779]
11. Angell M. Industry-sponsored clinical research: A broken system. JAMA 2008;300(9):1069-1071. [http://dx.doi.org/10.1001/jama.300.9.1069]
12. Steinbrook R. Gag clauses in clinical-trial agreements. N Engl J Med 2005;352(21):2160-2162. [http://dx.doi.org/10.1056/NEJMp048353]
13. Kjaergard LL, Als-Neilsen B. Association between competing interests and authors' conclusions: Epidemiological study of randomized clinical trials published in the BMJ. BMJ 2002;325(7358):249-252. [http://dx.doi.org/10.1136/bmj.325.7358.249]
14. Als-Neilsen B, Wendong C, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials. JAMA 2003;290(7):921-928. [http://dx.doi.org/10.1001/jama.290.7.921]
15. Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005;294(2):218-228. [http://dx.doi.org/10.1001/jama.294.2.298]
16. Pereira TV, Horowitz RI, Ioannidis JPA. Empirical evaluation of very large treatment effects of medical interventions. JAMA 2012;308(16):1676-1684. [http://dx.doi.org/10.1001/jama.2012.13444]
17. Editorial. Clinical practice guidelines and conflict of interest. CMAJ 2005;173(11):1297.
18. Ross JS, Hill KP, Egilman DS, Krumholz HM. Guest authorship and ghostwriting in publications related to rofecoxib. JAMA 2008;299(15):1800-1812. [http://dx.doi.org/10.1001/jama.299.15.1800]
19. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS ONE 2009;(5):e5738. [http://dx.doi.org/10.1371/journal.pone.0005738]
20. Hall JC. How to dissect surgical journals: VIII – Comparing outcomes. ANZ J Surg 2011;81(3):190-196. [http://dx.doi.org/10.1111/j.1445-2197.2010.05358.x]
21. Luus HG, Muller FO, Meyer BH. Statistical significance versus clinical relevance. S Afr Med J 1989;76:568-570.
22. Hadler NM. The Last Well Person. Montreal and Kingston: McGill-Queen's University Press, 2004:35-43.
23. Davidson RA. Does it work or not?: Clinical versus statistical significance. Chest 1994;106(3):932-934. [http://dx.doi.org/10.1378/chest.106.3.932]
24. Devereaux P, Yang H, Yusut S, et al. Effects of extended release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): A randomized controlled trial. Lancet 2008;371(9627):1839-1847. [http://dx.doi.org/10.1016/S0140-6736(08)60601-7]
25. Deans KJ, Minneci PC, Danner RL, et al. Practice misalignments in randomised clinical trials: Identification, impact, and potential solutions. Anesth Analg 2010;111(2):444-450. [http://dx.doi.org/10.1213/ANE.0b013e3181e63976]
26. Vincent J-L. We should abandon randomized controlled trials in the intensive care unit. Crit Care Med 2010;38 (10):S534-S538. [http://dx.doi.org/10.1097/CCM.0b013e3181f208ac]
27. Holmes D, Murray SJ, Perron A, Rail G. Deconstructing the evidence-based discourse in health sciences: Truth, power and fascism. Int J Evid Based Health 2006;4(3):180-186. [http://dx.doi.org/10.1111/j.1479-6988.2006.00041.x]
28. Osler W. On the educational value of the medical society. Address to the New Haven Medical Association, 6 January 2003. Yale Medical Journal 1903;ix:327.
29. Silverman ME, Murray TJ, Bryan CS, eds. The Quotable Osler. Philadelphia: ACP Press, 2003.

Accepted 20 May 2013.