

Editorial

Inappropriate, Exclusive Reliance on Objectivity

Raymond Maurice Kirk

University Department of Surgery, Royal Free Hospital, Pond Street, London, UK

Kuwait Medical Journal 2003, 35 (2): 83-85

"Not everything that can be counted counts, and not everything that counts can be counted"

Sign on the wall of Albert Einstein's office at Princeton University.

Many exciting changes are taking place in all areas of medicine. Much of it is commendable. We can improve only by searching for better ways of diagnosing and treating patients, developing methods of selecting the best aspiring doctors, and training, examining and appraising those we have chosen. It is important though, to assess proposed changes and not accept them merely because of their novelty.

OUR PATIENTS

We are pressed to practise 'evidence-based' medicine. The evidence stems from highly regarded and often specialized units. Based on their results 'guidelines,' 'algorithms,' 'protocols,' and 'flowcharts' are formulated. There are many variables in the aggressiveness and extent of diseases, the physical and mental state of patients, the choice of treatments and the facilities for delivering them. In order to limit the number of variables when comparing treatments, the investigators need to reject many patients who do not fulfil the requirements of the trial. The patients who present themselves to us, often do not conform to those of patients entered into trials and the guidelines do not strictly apply to them. If our management deviates from those laid down, and the outcome is poor, we may feel vulnerable to accusations of malpractice. As often happens, new methods are adopted over-enthusiastically and we need to ensure that guidelines are used as aids, not as firm rules of management.

The analysis of findings and results is usually carried out using statistical methods. To achieve this, continuous variables must be converted into numbers – analogue to digital. Analogue scales are like shades of grey and often subjective, for

example, determining the severity of post-operative pain. Patients may be asked to identify the level of pain they experience following a procedure, on a scale of 1 to 10. The resulting figures are then used for statistical analysis and have a specious appearance of objectivity. Digital scales are often black and white, yes or no, survival or death. Do not automatically accept numbers and statistics as scientifically valid. A journalist, Malcolm Muggeridge, stated, 'The credulity which western man accords to any numerate answer would be the envy of any African witch doctor.'

There is insufficient time to read every article that appears and we tend to accept the findings if the authors are distinguished, the institution where they work is well known and the journal is prestigious. We assume that the referees have checked the methods, interpretation of the findings and validity of the conclusions. Do not take this for granted. It is a valuable exercise to read an occasional paper of interest to you while questioning every statement. A surprising number are not clearly defined or justified, for example, how exactly were patients followed up, by whom, for how long, and how often? The longest follow-up is often given but not the shortest. Be suspicious of long lists of references and wonder if the articles quoted have been read by the authors or merely listed from the many sources now available, in the hope of adding authority to their claims.

Review articles tempt us to avoid careful, critical reading of individual reports. Remember, though, that reviewers, who act as inquisitors, already have their own point of view. They may unconsciously select, from the wealth of publications, references that support their viewpoint and those opposing opinions that can be refuted. A better balance is the adversarial style adopted in some American journals in which invited commentators point out weaknesses in the arguments, and propose alternative or opposing interpretations or methods.

Address correspondence to:

R M Kirk, Univ. Department of Surgery, 9th Floor, Royal Free Hospital, Pond Street, London NW3 2QG, UK. Tel: 0044 20 7794 0500 Ext 4452, (R) 0044 20 8340 8575. Fax: 0044 20 7472 6444 - E-mail: kirk@rfc.ucl.ac.uk

The most highly regarded method of comparing possible courses of action is randomised, controlled, prospective, double blind trials, where this can be applied. Very few investigators start from a completely neutral standpoint. They usually wish to demonstrate that their favoured methods are superior. Only when others challenge the conclusions and repeat the findings does a balanced view emerge.

OUR PROFESSION

Medicine, our wonderful profession, offers a niche for every type of personality, every type of aspiration. There is increasing pressure towards standardizing the assessment and selection procedures for medical students, and also the examination of candidates and applicants for occupational appointment. Those driving the changes must supply evidence on which to change practice.

Educationalists claim that objective assessments give equal opportunity and exclude subjective bias. In many cases each interviewer must put the same questions, after approval by the chairman, to all the candidates. Because candidates differ, the permitted questions are rarely appropriate for all of them. At the subsequent assessors' discussion the reasons for preferring or rejecting each candidate must be 'objective,' but all the evidence is not objective. Two candidates may give similar replies to a question but their manner may differ. We may trust one reply and distrust the other but if we express this opinion we may be accused of discrimination.

What the educationalists cannot demonstrate is that these assessments test the aptitudes necessary for the successful practice of clinical medicine. No one has satisfactorily compared the results in these, or any other tests, with the careers of the candidates. Is success in clinical medicine the acquisition of a departmental professorial chair, a valuable addition to research, a lucrative private practice, or a sound, caring general practitioner working in a deprived community? They all need individual drives, commitment, knowledge base and indefinable skills. Skills are often disparaged in comparison with theoretical knowledge, as being the qualities of an artisan rather than of a professional. It has been pointed out by Michael Polanyi that skill is a separate and valuable form of wisdom – and not to be acquired merely from books^[1].

Favoured examinations that are claimed to be objective include multiple choice questions and extended matching tests because they can be marked by computers. However, they do not require the candidate to supply the correct answer,

only to identify it. Does it represent the thought processes of a clinician, or merely check a fact^[2]? Objective structured clinical assessments are also favoured – but do they really test the complex thought processes we need to reach a diagnosis and treat our patients^[3]? Facts can be looked up. It can be argued that in the USA no test of clinical skills is carried out before licensing a physician to practice, since the 1960s^[4]. The high quality of most American doctors suggests it is unnecessary. Does a single encounter at a clinical examination offer a satisfactory test of these skills?

Aptitude is not yet reliably testable but is more important than factual knowledge, most of which can be found in books. It has been attempted by the Association of Surgeons of the Netherlands. In the hope of selecting the most appropriate candidates for training, they recruited the help of industrial psychologists to identify the desirable characteristics and produce tests for each of them^[5]. Of course, none of us is endowed with all the aptitudes; we compensate, or overcompensate for our weaknesses. Moreover, the Gestalt psychologists in Germany pointed out at around the turn of the 20th century, that we cannot fully judge the whole by examining the parts. The full character of a person emerges only after close observation over extended contact. This has traditionally been accomplished through the master/apprentice relationship.

TRUST

Patients and colleagues need to trust us. Reputations are earned, not acquired by training and certification, gaining the title of 'Doctor,' or adopting a uniform. They are not objectively assessable – indeed we cannot always describe why we trust one person but not another. We walk into a room and for reasons that we cannot yet identify, we have faith that some are worthy of our trust and we are suspicious of others.

Within the master-apprenticeship relationship, monitored increasing responsibility can be metered out until the master considers the trainee capable of functioning autonomously. The most important question a clinical trainer, examiner, member of appointment committees, has to answer is, 'If this person is on call for me at night, shall I sleep well?' The perceived attitude of the candidate is more important than the spoken or written words. It is not sufficient to test examination candidates and applicants for appointment to determine if they can perform. A much more fundamental question is, 'Will they perform?'

In 2002 Onora O'Neill broadcast on the British Broadcasting Corporation radio, a series of 5 lectures named in honour of the first Director General, Lord Reith, entitled, 'A question of trust.'

She challenged the topical obsession with transparency of information, arguing that as we obsessively examine the behaviour of others we erode our trust in them. One of the major characteristics that facilitate good practice is loyalty and this is generated by mutual trust. Professor O'Neill argues that intrusive, obsessive methods of assuring accountability and transparency are counter-productive, damaging, rather than reinforcing trustworthiness.

Doctors are sometimes compared with airline pilots who are regularly appraised. There is increasing pressure for the results of treatment in hospitals, and even by individual doctors, to be ranked and published. The practice of medicine is not parallel with flying an aeroplane. The function of even very complex machines, such as aeroplanes, can be measured. A pilot can refuse to fly an aeroplane that has a fault, or to fly if there are external dangers. A doctor does not manage a machine but a complex biological human being with many unmeasurable, often unidentifiable characteristics. Our patients come because they have an imperfection and we may need to treat them, even when circumstances and facilities are defective.

Airline pilots have the duty to report any defects they encounter and errors they make^[6,7]. They are not punished for their errors, only for failing to report them. It is accepted that all humans make errors, and that if an error can be made, sooner or later it will be made. The subsequent enquiry investigates how the system can be changed to

protect against the same error being repeated. Although doctors should behave similarly, they may fear that if they report an error they will be blamed and no consideration will be given to the circumstances - and the possibility of revising them to protect against a recurrence.

Professor O'Neill concluded that the real enemy of trust is deceit. Doctors need to retain the trust of our patients, our colleagues and of the public in general. To do this we must be able to declare problems and errors openly, confident that our commitment to good practice is not doubted.

REFERENCES

1. Polanyi Michael. Personal knowledge. Routledge Keegan and Paul. London 1973; 49-65
2. Fleming PR, Sanderson PH, Stokes JF, Walton HJ. Examinations in Medicine, Churchill Livingstone, Edinburgh 1976 p vii
3. Cox K. No Oscar for OSCA. Medical Education. 1990; 24: 540- 545
4. Gough MH, Holdsworth R, Bell JA, Keenan JN, Lagaay MB, Ven de Loo RPJM, Droog A. Personality assessment techniques and aptitude testing as aids to the selection of surgical trainees. Annals of the Royal College of Surgeons of England 1988; 70: 265-279
5. Winchester DE, Ruscher-Rogers K. A clinical examination for medical students? Letter. New England Journal of Medicine 2003; 348:1294-1295.
6. Reason J. Human error. New York: Cambridge University Press; 1990.
7. Shappell S, Wiegmann D. Applying Reason: The human factors analysis and classification system. Human Factors and Aerospace Safety 2001; 1:59-86.