

subjective. Lord Lister was observed to have “none of the dramatic dash and haste of the surgeon of previous times ... he proceeded calmly, deliberately, and carefully.” As he told his students, “Anaesthetics have abolished the need for operative speed and they allow time for careful procedure.”<sup>5</sup> Junior surgeons have been ranked using global scores based on subjective criteria, but multiple observers are needed to obtain acceptable reliability.<sup>9</sup> Gathering panels of experts to watch videos or attend theatre may be possible for a research project but is expensive in manpower and time, and this limits its feasibility in real life.

Assessment may be easier in the surgical skills training laboratory than in theatre. Surgeons may behave differently under simulated conditions, but if the tasks are designed carefully to reflect real surgical practice such tests could fulfil the essential requirements of feasibility, reliability, and validity.<sup>10</sup> Abstract tests of manual dexterity have not stood up to validation<sup>11</sup> and would appear to be so far removed from the act of surgery as to be unhelpful in selecting potential surgeons. Subjective methods using structured scoring systems have been shown to be reliable.<sup>12</sup> Although multiple observers were used to rate candidates in terms of “economy” and “fluidity” of movement, it was difficult to validate these scores with subjective rankings of residents in the operating theatre.<sup>13</sup>

Recent work has tracked the movement of laparoscopic surgical instruments in the laboratory. Objective measurements of economy of motion and number of movements made are generated by the assessment device. These criteria have been validated for tasks in both reality and virtual reality.<sup>14 15</sup> Devices that objectively and reliably quantify surgical dexterity could have advantages over traditional subjective evaluation, particularly as a screening tool.

A system that can provide unbiased and objective measurement of surgical precision (rather than just speed) could help training, complement knowledge based examinations, and provide a benchmark for certification. A specific and sensitive test of operative com-

petence could also detect an important problems and might improve surgical outcome. Revealing underperformance early would allow for further training or career guidance towards other less practical specialties. The surgical profession needs a reliable and valid method of assessing the operative skill of its members. A driving test may not be a guarantee against accidents but it makes it less likely that you career off the road. Surgeons, the public, and politicians need reassurance.

Ara Darzi *Professor of surgery*  
 Simon Smith *Clinical research fellow*  
 Nick Taffinder *Specialist registrar*

Division of Surgery, Anaesthesia, and Intensive Care, Imperial College School of Medicine at St Mary's, London W2 1NY (a.darzi@ic.ac.uk)

- 1 Poloniecki J, Valencia O, Littlejohns P. Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery. *BMJ* 1998;316:1697-1700.
- 2 Spencer F. Teaching and measuring surgical techniques: the technical evaluation of competence. *Bull Amer Coll Surg* 1978;63:9-12.
- 3 Fielding LP, Stewart-Brown S, Dudley HA. Surgeon-related variables and the clinical trial. *Lancet* 1978;ii:778-9.
- 4 Smith R. All changed, changed utterly. *BMJ* 1998;316:1917-8.
- 5 Ellis H. *Famous Operations*. Media, Pennsylvania: Harwal, 1984.
- 6 Van-Rij AM, McDonald JR, Pettigrew RA, Putterill MJ, Reddy CK, Wright JJAD. Cusum as an aid to early assessment of the surgical trainee. *Br J Surg* 1995;82:1500-3.
- 7 Hanna GB, Shimi SM, Cuschieri A. Randomised study of influence of two-dimensional versus three-dimensional imaging on performance of laparoscopic cholecystectomy. *Lancet* 1998;351:248-51.
- 8 European Association of Endoscopic Surgeons. Training and assessment of competence. *Surg Endoscopy* 1994;8:721-2.
- 9 Winkel CP, Reznick RK, Cohen R, Taylor B. Reliability and construct validity of a structured technical skills assessment form. *Am J Surg* 1994; 167:423-7.
- 10 Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993; 165:358-61.
- 11 Graham K, Deary I. A role for aptitude testing in surgery? *J R Coll Surg Edinb* 1991;36:70-4.
- 12 Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84:273-8.
- 13 Faulkner H, Regehr G, Martin J, Reznick R. Validation of an objective structured assessment of technical skill for surgical residents. *Acad Med* 1996;71:1363-5.
- 14 Taffinder N, Smith S, Jansen J, Ardehali B, Russell R, Darzi A. Objective measurement of surgical dexterity - validation of the Imperial College Surgical Assessment Device (ICSAD). *Minimally Invasive Therapy and Allied Techniques* 1998;7(suppl 1):11.
- 15 Taffinder N, McManus I, Jansen J, Russell R, Darzi A. An objective assessment of surgeons' psychomotor skills: validation of the MIST-VR laparoscopic simulator. *Br J Surg* 1998;85(suppl 1):75.

## Pleasing both authors and readers

*A combination of short print articles and longer electronic ones may help us do this*

To succeed, journals need to please both authors and readers. There is, however, a tension between the needs of the two, particularly when the authors are mostly researchers and the readers mostly practitioners. Practitioners like research articles to be short and sweet, whereas researchers want—rightly—to include enough material for critical readers (often other researchers) to be able to appraise the study and if necessary repeat it and also, increasingly, to be able to include it in a systematic review. Journals have struggled with this tension for years, and often the result is that we please nobody. Research among readers consistently shows that research articles are not well read, while many studies have shown that essential data are often missing from

research reports. Now the electronic revolution offers us a chance to please both readers and authors simultaneously.

Today's *BMJ* includes four papers where a short version is published in the paper journal and a longer version in the electronic journal (*eBMJ*) (p 897-914).<sup>1-4</sup> We even have an acronym for the process: ELPS (electronic long, paper short). This first effort is an experiment, and we are not yet planning to introduce this system for every research article—but we may if both readers and authors are pleased.

Our experiment follows an intense debate within the journal on whether this is a good idea. The arguments in favour are those we've already advanced plus the possibility of using the pages spared for sections that are

*Papers pp 897-914*

*BMJ* 1999;318:888-9

more popular than research articles. The main arguments against are that the paper journal is the "proper" journal and that not everyone has access to the internet. However, we have already said that the *eBMJ* is the primary journal in that it includes everything published in the paper journal and an increasing amount more.<sup>5</sup> Paper and electronic versions of journals will diverge as the electronic versions exploit the full potential of the internet. The journal *Pediatrics*, for example, publishes some studies primarily in electronic form, with only an abstract in the paper version.

The other major concern is that some people, particularly in the developing world, do not have easy access to the internet. In the developed world access to the internet is increasing exponentially, and soon it will be accessible through television without any need for a computer. Most researchers have easy access through academic networks, and those who do not have direct access can easily obtain a copy of an article from the *eBMJ* in the traditional way through their medical library. In fact access to the *eBMJ* is probably easier than access to a print copy for anyone who is not a subscriber: since the *eBMJ* is free any library or other institution with an internet connection can access it immediately. Sadly, access to paper journals has been severely restricted in the developing world, and in the long run electronic forms of journals are likely to reach many more people than paper forms ever could. If we need at some stage to charge for the *eBMJ* then we will keep it free to those in the developing world (which we can at no extra cost to us, whereas the cost of transporting paper is substantial).

A further argument against ELPS is that it may become a licence for authors to produce interminable verbose reports. This we will resist, although reports on scientific studies may eventually expand to include sound, video, original data, software, and more. The challenge is not just to present studies in the same old way but to find ways to use the medium to full scientific advantage for both authors and readers/viewers. We accept that we could do a better job of including more essential information in paper versions of studies with-

out necessarily making them any longer. Standard formats—such as CONSORT for publishing reports of randomised controlled trials<sup>6</sup>—should increase the informative value of articles, though they do generally seem to make reports longer.

As ELPS is currently experimental, we are keen to present papers in different ways. The shorter versions in this week's journal have been made shorter by general shortening throughout all the sections of the paper, and in one case we have prepared two shorter versions, one much shorter than the other (pp 908, 912).<sup>4</sup> We have, however, debated whether to increase the readability of reports by emphasising the introduction and discussion or whether to help those readers interested in critically appraising studies by concentrating on methods and results. We hope to continue the experiment by presenting further different sorts of shorter versions. We look forward to hearing views on the whole idea and on how we might best use paper and electronic media to complement each other.

Tony Delamothe *Web editor*, BMJ

Marcus Müllner *Editorial registrar*, BMJ

Richard Smith *Editor*, BMJ

- 1 Williams S, Poulton R. Twins and maternal smoking: ordeals for the fetal origins hypothesis [abridged version]. *BMJ* 1999;318:897-900. (full version: [www.bmj.com/cgi/content/full/318/7188/897](http://www.bmj.com/cgi/content/full/318/7188/897))
- 2 Bredin M, Corner J, Krishnasamy M, Plant H, Bailey C, A'Hern R. Multi-centre randomised controlled trial of nursing intervention for breathlessness in patients with lung cancer [abridged version]. *BMJ* 1999;318:901-4. (full version: [www.bmj.com/cgi/content/full/318/7188/901](http://www.bmj.com/cgi/content/full/318/7188/901))
- 3 Quinn M, Babb P, Jones J, Allen E on behalf of the United Kingdom Association of Cancer Registries. Effect of screening on incidence of and mortality from cancer of the cervix in England: evaluation based on routinely collected statistics [abridged version]. *BMJ* 1999;318:904-8. (full version: [www.bmj.com/cgi/content/full/318/7188/904](http://www.bmj.com/cgi/content/full/318/7188/904))
- 4 Whitehead M, Drever F. Narrowing social inequalities in health? Analysis of trends in mortality among babies of lone mothers [abridged version]. *BMJ* 1999;318:908-14. (full version: [www.bmj.com/cgi/content/full/318/7188/908](http://www.bmj.com/cgi/content/full/318/7188/908))
- 5 Delamothe T, Smith R. The BMJ's website scales up. *BMJ* 1998;316:1109-10.
- 6 Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomised controlled trials: the CONSORT statement. *JAMA* 1996;276:637-9.

## Measuring the performance of public health agencies

*Government, like doctors and hospitals, should meet quality standards*

Performance measurement is a first step towards quality improvement in health care. When systems are in place to measure performance we can reward good performance, develop and evaluate ways of improving performance, and certify (or decertify) providers who perform (or don't perform) according to established standards. In the United States an extensive machinery has emerged to measure and ensure performance, but so far it has not been applied systematically to public health agencies. That situation is about to change, and, as it does, it opens up the interesting possibility of holding elected politicians even more accountable for decisions that affect health.

Efforts to measure the performance of healthcare providers in the US have expanded rapidly over recent

years. Much of this activity has been driven by the main purchasers of health care—large employers and government.<sup>1</sup> The Joint Commission on Accreditation of Healthcare Organizations accredits 18 000 facilities in the US, including hospitals, home care agencies, long term care facilities, and clinical laboratories, allowing them to participate in the federal Medicare programme for the elderly.<sup>2</sup> The performance of health maintenance organisations is assessed through "report cards"—especially the health plan employer data and information set<sup>3</sup>—and through accreditation by the National Committee for Quality Assurance.<sup>4</sup> Doctors must pass examinations to be licensed to practise medicine, and board certification in their specialties is often required to join medical group practices, to care for members of health maintenance organisations, and