Editorial II

Assessment of clinical competence

The most common way of assessing competence at the undergraduate and postgraduate levels in the UK has been the traditional clinical viva. In an era where the quality of assessments, in terms of validity and reliability, is under increasing scrutiny, this approach has come under mounting criticism. Evidence demonstrating the unreliability of the traditional clinical examination was published in North America in the 1960s but seems to have had little impact on assessment practices in the UK and countries with similar medical education and training systems. This can partly be explained by the lack of any obvious alternatives.

The last 20 years have seen some dramatic changes in research and development in methods of assessment. For example, the advent of the Objective Structured Clinical Examination (OSCE) has provided a viable alternative to the traditional approach and has become increasingly popular, particularly at the undergraduate level. The OSCE has mutated into many versions, ranging from the classic form of multiple short (5–10 min) stations involving examiners rating candidates on a structured marking sheet, through to multiple long stations (30 min) involving candidates undertaking a full clinical assessment on simulated patients trained to score the performance. An example of the latter is the new Clinical Skills Assessment being used by the Educational Commission for Foreign Medical Graduates in the USA.

Research on multistation examinations has produced valuable technical data which provide us with evidence about how to design such tests to achieve high levels of reliability and guidelines on how to construct them to demonstrate acceptable validity. However, valuable as they are for certification purposes, they do not provide a panacea and some important components of clinical competence are not readily or even appropriately assessed under examination conditions. These include attributes in many areas involving in-depth interaction between the student or doctor and patients, relatives, colleagues or others and where attitudes and ethical considerations play a major part in determining performance. Such areas inevitably require assessment in more complex situations and over an extended period of time in the workplace. Unfortunately, the methodology available to undertake assessment of these vital components is far less well developed and likely to be prone to problems of unreliability, perhaps even more so than the traditional clinical examination.

The article by Greaves and Grant in this issue examines some of the matters relating to the in-service assessment of anaesthetic trainees. The authors point out that measures of clinical outcome are not suitable for this task as they are often influenced by many factors unrelated to the performance of the doctor. They believe that the well accepted approach of observing clinical practice must be an important facet of the assessment of competence. They set out to investigate how this is currently undertaken and how it might be made more standardized without losing the flexibility of the present approach to evaluate the integration of technical, professional and judgemental skills into an overall performance. The results are the outcome of focus group discussions involving a panel of consultant and trainee anaesthetists. They identified the approaches used by supervisors of research and development in methods of assessment. For example, the advent of the Objective Structured training (interrogation, documentary evidence, observation). However, how they organized the information and used it to judge performance seemed less clear. Nevertheless, consultants seemed confident that the assessment they made was valid and reliable. Unfortunately, no evidence that this is the case is provided and past experience would suggest that this confidence might not be well founded. On the other hand, the consultants did recognize the value of multiple observations by different observers as a requirement to enhance reliability and fairness. What was interesting was a recognition that certain patterns of behaviour might be more critical in recognizing incompetence; these included lack of flexibility, failure to adapt to changing circumstances, irrational judgements and a disordered approach. The outcome of the study included a definition of attributes of competent practice with associated statements of appropriate and inappropriate behaviours. This was linked to a more systematic strategy for assessment than would currently be the case. The authors conclude by pointing out that such a process needs to be tested on large numbers of trainees in order to establish levels of validity and reliability.

Such studies are an important first step in the development of improved methods of performance-based rather than competence-based assessments. The requirements of the recent changes to specialist training and the advent of ‘revalidation’ highlight the need for more defensible approaches to assessing performance in practice. Currently only limited research is available to guide progress but experience from competence-based testing alerts us to the fact that solutions will be more complex and more expensive (in time). We will have to rely on the judgement of peers making direct observations, a process that will enhance validity if appropriately
structured and targeted but which will inevitably cause problems with reliability. The latter concern will require the development of tools to guide the process and record the outcome as well as necessitate the training of assessors. However, these will have limited impact and, for high-stakes summative purposes, it will be necessary to combine the assessments from multiple observers over many instances or periods of practice.

D. I. Newble
Department of Medical Education
University of Sheffield
Coleridge House
Northern General Hospital
Sheffield S5 7AU
UK

References
6 Greaves JD, Grant J. Watching anaesthetists work: using the consultant’s professional judgement to assess the trainee’s developing clinical competence. Br J Anaesth 2000; 84: 525–33