Assessing the performance of specialist registrars

Andrew Wragg, Winnie Wade, Geraint Fuller, George Cowan and Peter Mills

ABSTRACT – Assessing the performance of doctors while they are engaged in clinical work is a challenging concept. The introduction of objective-based curricula provides the stimulus and opportunity for the Royal Colleges of Physicians to develop relevant and reliable methods of in-service assessment. We propose to pilot a study investigating the validity, reliability and feasibility of three assessment methods – direct observation of the clinical encounter using an adapted mini-CEX, direct observation of the performance of practical procedures (DOPS), and the doctor's ability to perform effectively as part of a team using 360° assessment. The methods will be studied in the setting of routine clinical care. Whilst demanding of time from both trainees and trainers, they will represent a significant advance on the current system which is characterised by a lack of evidence in the assessment process.

KEY WORDS: assessment, competence, curriculum, performance, specialist registrar, standards

Introduction

Assessment of the performance of doctors has recently become an important issue both publicly, as a result of high profile cases such as the Bristol Heart case, and within the profession, following the redesign of higher specialist training. The assessment of doctors in an honest and objective manner is laid out as a fundamental part of practice in the General Medical Council's guidance, *Good Medical Practice*.[1]

At present, the assessment of specialist registrars (SpRs) in medicine is a subjective process. It is based on an educational supervisor 'signing up' a trainee as competent in a specific task in the Record of Training Book. This usually occurs after an arbitrary period of time has been spent by the SpR working in that area and often having completed an arbitrary number of procedures in the case of practical skills. There is no process by which the educational supervisor collects objective evidence to inform judgements about a trainee's competence. It has been acknowledged that there needs to be a greater emphasis on performance-based assessment. There are still no robust mechanisms for formal assessment, and poor performance is not reliably recognised or addressed. There is a real need to develop and implement reliable objective methods of assessment to form an integral part of training of SpRs.

The role of the new curricula

The Joint Committee of Higher Medical Training (JCHMT), which represents all three medical Royal Colleges in the UK, has for some time recognised that there needs to be a review of the approach to specialist physician training. As a result, the JCHMT and specialist advisory committees (SACs) of the medical specialities have rewritten the specialty curricula in line with current educational thinking. The new curricula, which were launched in December 2002, are now objective-based and encompass modern educational principles.

Although the content of the new curricula has not changed dramatically, the emphasis of it has. The curricula lay out clear objectives that trainees will have to achieve during their training programme set against clear standards of competence. These competencies specify that trainees should be able to carry out specialist assessment and treatment of their patients to consultant level.

Each curriculum is based on a template that includes:

- background
- aims
- objectives
- subject matter (divided into knowledge, skills and attitudes)
- teaching and learning methods
- assessment methods.

It is the changes in assessment methods that are likely to have the biggest impact on trainees and trainers.

Key Points

At present, the assessment of specialist registrars in medicine is a subjective process

The Royal Colleges of Physicians are piloting three methods of assessment of performance

Methods have been selected which relate directly to the new specialist curricula

Andrew Wragg MRCP, Educational Research Fellow*

Winnie Wade BSc PGCE MA, Head of Education*

Geraint Fuller FRCP, Consultant Neurologist, Gloucester Royal Hospital

George Cowan FRCP, Medical Director, JCHMT*

Peter Mills FRCP, Consultant Cardiologist, The London Chest Hospital

*Royal College of Physicians

Clin Med 2003;**3**:131–4

What is the purpose of assessment?

Whilst identifying the 'trainee in difficulty' is one purpose of an assessment, documenting that a trainee has achieved the standard necessary for gaining a Certificate of Completion of Specialist Training (CCST) is obviously extremely important. Additional purposes include guiding and encouraging professional development, identifying areas for improvement (of both SpRs and training programmes) and providing robust validation of performance for concerned stakeholders such as patients, society and SpRs themselves.[2]

For most trainees, assessment will be an opportunity for personal development and will be relatively non-threatening and positive. On the other hand, assessment to identify the poorly performing doctor carries potentially serious consequences and therefore must be objective, reliable and valid.

What should be measured when assessing specialist registrars?

The attributes of a good doctor are multiple and no single statement or phrase describes the perfect doctor. Attributes include generic skills and attitudes, such as communication and teamworking, as well as the specialised knowledge and skills relevant to each specialty. It is therefore important that it is clear what we are actually trying to measure and assess. The assessment of SpRs in the real world of medical practice is a complex process. Doctors work in teams and provide only a portion of the direct care that the patient receives within the healthcare environment.[3] It is the measurement of actual *clinical performance* that we must aim to assess.

What is meant by performance and how can it be assessed?

The pyramid of competence (Fig 1)[4] is a simple conceptual model that identifies the essential components of clinical competence: 'Knows>Knows how>Shows how>Does'. It is the assessment of 'Does' that is most relevant to a doctor's actual performance. Multiple choice questions, essays and oral examinations can test factual and applied knowledge but are inadequate in assessing clinical performance which includes skills and attitudes and the interaction of an individual doctor with his/her colleagues and patients. Although objective structured clinical examinations (OSCEs), long cases and short cases, and simulated patients are used to assess clinical skills, they occur in an artificial environment and hence lose some validity. They measure competence ('Shows how') rather than real clinical performance ('Does').

In addition, the organisation and delivery of formal assessments, eg OSCEs, have significant implications for resources and time. Ideally, the assessment process should measure the actual clinical performance in practice in a reliable and valid manner within an acceptable period of time. The description of real clinical performance should include assessment of the application of knowledge, and use of appropriate skills and attitudes which together make up clinical judgement. There is no single tool that can assess all these components but methods that are based on the direct observation of performance should provide a valid method for assessing clinical practice.[2]

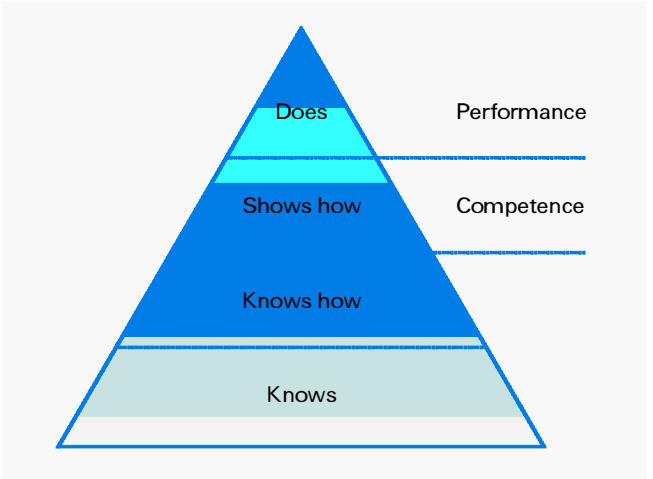

Fig 1. Miller's Triangle: the pyramid of confidence.[4]

What assessment tools should be used?

'The goal of assessment in medical education remains the development of reliable measurements of performance which, as well as having predictive value for subsequent clinical performance also have a formative, educational role.'[5] Because of the variability in scoring with almost all methods of assessment, each assessment method has to be repeated on a number of occasions to ensure reliability and create as complete a picture as possible. The variability in the marking of assessors is a major source of error. In nearly all methods, an individual assessment may tell you more about the assessor than the trainee. Hence the need for multiple assessments by different individuals on different occasions, regardless of the methods used.

Sound assessment should be based on the following five principles:[2]

1 *Validity*: does it measure what it purports to measure?
2 *Reliability*: is the measure of performance consistent?
3 *Feasibility*: is the method practical in a busy clinical environment? What is the cost?
4 *Educational impact*: will the trainee learn from the process?
5 *Acceptability*: do the people involved have confidence in the process?

Since the consequences of assessment may have important implications for poorly performing SpRs, the procedure has to be founded on evidence-based medical educational principles. Applying the above five principles to potential methods of assessment should ensure that the parties involved have confidence in them and that the assessment methods would be defensible if their validity was questioned.[6]

Unfortunately, there is no known single assessment method that can adequately cover all five of these qualities and address

all of the complex abilities that a doctor requires. A combination of assessment methods is necessary.

The pilot of methods of assessment for specialist registrars

Considerable progress has already been made. The Royal Colleges of Physicians working with all of the SACs have now developed the new curricula and are exploring assessment methods. Methods have been selected which focus on the specific attributes laid down in the curricula and include the five principles of good assessment mentioned above. The Royal Colleges of Physicians are now running a pilot study to look at new methods of assessment following a systematically designed protocol. This pilot is taking place across the UK and involves SpRs and consultants from all medical specialties.

Three methods of assessment have been selected for the initial pilot to address important areas of real clinical performance: clinical skills, practical skills and generic skills.

The three methods that are being explored are:

1 Mini-CEX
2 Directly observed assessment of practical skills
3 360° assessments.

Mini-CEX

The Mini-CEX (clinical evaluation exercise) was developed in the USA to assess the clinical skills that trainees most often use in real clinical encounters. It was adapted from the traditional 'long-case' assessment. The Mini-CEX involves direct observation by an educational supervisor of a trainee's performance in real clinical situations and is designed to assess skills such as history taking, clinical examination, communication, diagnosis and management of patients and their problems. Direct observation appears to have high face validity (ie appears to measure what it purports to measure) but it is resource intensive. In addition, a trainee's performance under observation may well be different from his/her normal behaviour and hence it may be more of a measure of maximum competence rather than performance. The Mini-CEX should therefore be repeated on multiple occasions with different patients in different clinical environments. The Mini-CEX has been introduced and piloted in the USA and is now in widespread use for residency programs.[7] It is reproducible, offers trainees greater opportunity for actual observation and feedback and has proved to be a valid and reliable method of assessment. In addition, it provides increased opportunity for education because each trainee has a number of interactions with consultants and received feedback as part of each assessment.

Directly observed assessment of practical skills (DOPS)

It is essential that all trainees should be adequately assessed for competence in the practical procedures that they undertake. The traditional methods of assessment used are based on the log-book, including records of complications, and informal assessments by educational supervisors. This may be based on the time spent undertaking a specific procedure and the number of procedures performed. There is no substitute for quality experience but it is generally felt that a 'numbers-based' approach to the assessment of training is no longer sufficient. The numbers of procedures required to reach competence varies considerably between individual trainees and is also related to the quality of the trainer/training. A competence-based assessment is likely to be more valuable than a log of numbers.

DOPS is a method of assessment similar to the Mini-CEX developed specifically for the assessment of practical skills. The process requires an educational supervisor to observe directly the trainee undertaking the procedure being assessed. The assessor is required to make judgements about specific components of structured tasks, and then grade the performance. These include generic skills such as consent and communication, in addition to the details of the practical aspects of the procedure under assessment. The DOPS, like the Mini-CEX, is both formative and summative and the trainee receives constructive feedback at the end of each assessment. The DOPS has to be repeated on a number of occasions. The number of assessments required will be addressed by the pilot study.

The 360° assessment

The 360° assessment is an objective systematic collection and feedback of performance data, using a structured questionnaire, on an individual derived from a number of stakeholders on his/her performance. This assessment method has been used in industry and business to assess performance, to encourage improvements in employee performance and to inform decisions on promotion and job planning.[8] The American Board of Internal Medicine has also evaluated this instrument and found it to be a useful tool.[9,10] Essentially 'raters' are people with whom the individual works. In the case of SpRs, raters will include consultants, nurses, peers, secretaries and other allied health professionals. They will provide structured feedback about an individual's performance. The 360° assessment is a useful way of assessing generic skills such as communication, leadership, team-working, teaching, punctuality and reliability, accepting that they are an indirect measure of performance.[2] Although individual responses are of little value, multiple assessments can provide valuable information. This method of assessment has already been piloted with senior house officers in the UK and found to be both practical and acceptable to trainees in a district general hospital setting.[11]

What about an exit examination?

Formal knowledge-based assessments such as multiple choice questions may reliably test knowledge but they do not assess skills or attitudes. In addition, how trainees perform in examinations does not always correlate with how they perform in actual clinical practice.[12] Although a small number of the medical SACs are interested in formal knowledge-based assessments, an isolated

deficiency in knowledge is rarely a cause of major concern by the time a trainee has achieved an SpR appointment. Although knowledge-based assessment tools are likely to be developed, they will have a significant educational component as well as being used for assessment. They are likely to be case-based and available through the Internet.

Why is a pilot necessary?

Whilst the studies cited above suggest that these tools may prove to be helpful, it is not clear how they will work as assessment tools for SpRs in the UK. We cannot confidently extrapolate the use of the Mini-CEX in first-year residents in the USA (equivalent to SHOs in the UK) to SpRs in the UK. We cannot infer that the attributes of 360° assessment within the NHS will reflect those in business. These studies have to be repeated within the setting in which they are intended to be used, with content appropriately modified for use within the UK and the NHS. Because of the impact of introducing a more rigorous approach to assessment we must be certain of the reliability, validity, feasibility and acceptability of these new methods before using them in the real world.

What will be the benefits for the trainee?

Most trainees will achieve their CCST regardless of the changes in assessment. Having frequent supervised assessments with constructive feedback from a consultant can only be beneficial. With the increasing workloads of consultants, less time is available for training. These encounters, be they assessment of clinical or practical skills, will also provide valuable training opportunities for SpRs.

For the poorly performing trainee with difficulties, this process should allow problems to be identified at an early stage allowing remedial action to be taken. It is unhelpful for all involved if problems with an individual's performance are only identified at a penultimate year assessment or worse still when appointed as a new consultant. Having an effective and robust assessment will allow the trainee with difficulties to be identified at an early stage and have his/her training modified. Deficits in competence and performance are often as much the failing of the trainers or the programme as they are of the trainee.

Most people, including SpRs, are aware of individuals who may be lacking in the generic skills required of a doctor. The problem of 'inappropriate' attitudes has been formally recognised by the medical profession, both in the USA and in the UK.[13] These skills are not always given the same importance as technical skills. Because of the lack of objective evidence of these generic attributes, such as communication skills or teamworking, most trainers find it extremely difficult to address these difficult issues. Training programmes are letting trainees down when they acquire their CCST with poor generic skills. This not only makes it harder for them to find a consultant post but also ill equips them for professional life as a consultant.

Having robust documentation should also prove to be valuable for trainees. If an individual's competence were ever called into question in the future then sound documentation of assessment of performance would be of considerable value. In addition, having sound documentation of assessment will allow Record of In Training Assessment (RITA) panels to develop an accurate picture of the trainee.

How much time will this all take?

It is inevitable that any new methods will require more time on behalf of the trainee and trainers. The modern climate in which we work demands that we have a more objective assessment and the cost of this is the use of already limited time. However, the Royal Colleges of Physicians have developed and adapted methods of assessment that can be conducted as an integral part of the normal working day with the minimum impact on time. The assessment of performance of doctors in training must be a positive development for patients, consultants and trainees.

References

1 General Medical Council. *Good Medical Practice*, 3rd edition. London: GMC, 2001.

2 Hays RB, Davies HA, Beard JD, Caldon LFM *et al*. Selecting performance assessment methods for experienced physicians. *Med Educ* 2002;**36**:910–17.

3 Southgate L, Hays RB, Norcini J, Mulholland H *et al*. Setting performance standards for medical practice: a theoretical framework. *Med Educ* 2001;**35**:474–81.

4 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;**65**:563–67.

5 Wass V, Van derVluten C, Shatzer J, Jones R. Assessment of clinical competence. *The Lancet* 2001;**357**:945–9.

6 Schuwirth LWT, Southgate L, Page GG, Paget NS *et al*. When enough is enough: a conceptual basis for fair and defensible practice performance assessment. *Med Educ* 2002;**36**:925–30.

7 Norcini J, Blank L, Arnold G, Kimball H. The Mini-CEX: a preliminary investigation. *Ann Intern Med* 1995;**123**:795–9.

8 Waldman D, Atwater L. *The power of 360° feedback*. Houston, Texas: Gulf Publishing, 1998.

9 Ramsey P, Wenrich MD, Carline JD, Innui TS *et al*. Use of peer ratings to evaluate physician performance. *JAMA* 1993;**269** (13):1655–60.

10 Wenrich MD, Carline JD, Giles LM, Ramsey PG. Ratings of the performance of practicing internists by hospital-based registered nurses. *Acad Med* 1993;**68**(9):680–7.

11 Whitehouse A, Walzman M, Wall D. Pilot study of 360° assessment of personal skills to inform record of in-training assessments for senior house officers. *Hosp Med* 2002;**63**:172–5.

12 Rethans JJ, Norcini JJ, Baron-Maldonado M, Blackmore D *et al*. The relationships between competence and performance: implications for assessing practice performance. *Med Educ* 2002;**36**:901–9.

13 Martin J, Lloyd M, Singh S. Professional attitudes: can they be taught and assessed in medical education?. *Clin Med* 2002;**2**:217–23.