

# Lessons from the UK: Doctors' views of changes in postgraduate training

A. Sabey<sup>1</sup>, K. Feest<sup>2</sup> & S. Gray<sup>2</sup>

---

## Abstract

**Aims:** To explore the views and experiences of trainee doctors and their assessors undertaking competence assessments, in the first year of the new postgraduate programme in the UK.

**Methods:** A qualitative approach using individual, semi-structured interviews with seven first year trainees and seven assessors from across a range of specialties in a large acute hospital Trust in the UK.

**Results:** Assessing competence of newly qualified doctors has the potential to bring important benefits such as fostering relationships between junior and senior staff, building confidence in early days in practice and providing an early warning system for doctors who are struggling. However, certain barriers exist which make the benefits difficult to realise, and collectively undermine the value of assessment. Principally, inadequate preparation for the role of

assessor and lack of time for assessment are revealed, together with perceptions of bias and a lack of rigour in the use of assessment tools.

**Conclusions:** The role of competency assessment in postgraduate medical education is expanding and the findings of this UK study will be relevant to those implementing new systems as in Australia and New Zealand to avoid the problems seen here. The most significant is the importance of preparing staff for new roles in assessment including training assessors and recognising the training role in the form of protected time. This will require the full commitment of professional, regulatory and employing bodies. Further strategies to aid success include optimising the number of assessments and maximising objectivity by reducing choice of assessors.

**Keywords:** Competence assessment; medical education; postgraduate training; qualitative.

---

1 University of the West of England, Bristol, UK

2 Severn Deanery, Bristol, UK.

## Correspondence

Abigail Sabey  
Senior Lecturer  
Faculty of Health and Life Sciences  
University of the West of England  
Gloucester Campus  
Hartpury  
Gloucester GL19 3BE, England  
UK

---

## Introduction

Postgraduate training for medical graduates in the UK has undergone radical reform since the launch of the Government review *Modernising Medical Careers* in 2003 and the introduction of the Foundation Programme (FP) in August 2005. This saw for the first time, a national curriculum for junior doctors' training and structured, ongoing assessment of competence throughout a two year training period forming the 'bridge between undergraduate...and specialist and general practice training' (Department of Health 2004). Similar reforms are now facing Australia and New Zealand driven by similar imperatives of ageing populations and the resulting expansion in the medical workforce (Hays 2007), in turn triggering upheaval in postgraduate training and formal systems of performance assessment. (Fox and Arnold 2008). Adjustment to such significant change cannot be expected to happen quickly: in the UK the potential benefits from the new system are not yet fully realised (Gray et al 2009), making it all the more timely to consider the lessons that can be learned from the UK experience.

The assessment of competence of qualified physicians has been the focus of considerable international debate in the past 10 years and a number of accepted approaches have emerged. (Norcini et al 2003, Holmboe et al 2003, Holmboe et al 2004, Southgate et al 2001a, Violato, Lockyer and Fidler 2003). The UK FP drew upon this literature in developing four tools to assess the range of competencies required of new doctors (see Box 1). However, whilst the methods for assessment of competence

**Box 1.:** Outline of assessment tools used within Foundation Programme (see <http://www.hcat.nhs.uk/foundation/> for actual copies of the tools)

### Mini-Clinical Evaluation Exercise (Mini-CEX)

A 15-minute observation of a doctor/patient interaction, chosen by the trainee, designed to assess clinical skills, attitudes and behaviours of trainees. Six should be completed during the first year ('F1' year). The original mini-CEX was developed in the U.S. and has been shown to have high reliability (Norcini et al 2003) and validity (Holmboe et al 2003).

### Direct Observation of Procedural Skills (DOPS)

The observation of a range of practical procedures using a structured checklist to provide feedback. Six should be completed during F1. DOPS was developed by the Royal College of Physicians to improve the reliability and validity of such observation (Wilkinson, Benjamin and Wade 2003).

### Case-based Discussion (CbD)

A short, structured discussion between trainee and assessor of actual cases using patient's notes with the trainee's written entries. At least six are required during F1. A version of the CbD originated in the U.S. and has been used as part of the General Medical Council's performance procedures (Southgate et al 2001b) which informed the development of the CbD tool for the FP.

### Mini-Peer Assessment Tool (Mini-PAT)

Based on the domains of Good Medical Practice (General Medical Council 2001) this tool seeks feedback from a range of eight colleagues, nominated by the trainee, including the supervising consultant. Self-assessment is also undertaken. Two assessments are completed per year. The mini-PAT is closely based on a peer-review assessment tool developed in Sheffield, U.K. which has been shown to be both valid and reliable for use in the UK (Archer and Davies 2004) with a variety of grades of doctor (Beard et al 2005).

have been the subject of research and evaluation, little is known of the value and impact of assessment for trainees and supervisors involved. This study uses qualitative research methods to explore in depth, the views of trainees and their supervisors undertaking assessments in the first full year of the FP in the UK. From taking an in-depth approach with a typical group it aims to achieve understanding of the real demands arising from a complex new system of assessment and with it, a snapshot of the initial response to widespread change in postgraduate medical education.

## Methods

Data were collected via individual, semi-structured interviews with 14 participants (seven trainees; seven assessors). Participants were recruited from a large acute hospital NHS Trust in the South West of England, UK between May and August 2006. The Trust provides all medical and surgical services from two large acute hospitals, to a predominantly urban population of 500,000 with some spread into adjoining suburban and rural areas; and some specialist regional services reaching over 2 million people.

The final sample of trainees and assessors was achieved from a total of 54 who were invited to take part in the study from the first year of the new FP (F1). Initial sampling was purposive with selection of trainees and supervisors being guided by a clinical tutor with knowledge of potential participants, with network sampling being used later to boost participation. Interviews were completed with four male and three female trainees reflecting gender and the range of hospital specialties,

together with two female and five male assessors from the pool of assessors across the branches of medicine and surgery. Therefore some assessors were involved in the assessment of the trainees in the study although participants were not asked to disclose specific information of this kind. Six assessors were consultants with an educational role in the FP, and one was a nurse manager.

All interviews were carried out by the researcher (AS), based on a short schedule of topics derived from the literature. The five main topics raised were: general impressions of assessment and the four tools; practical aspects/organisation; educational value and validity; feedback; impact and outcomes of assessment. In keeping with the qualitative approach, it is important to reflect on and recognise the researcher as part of the process of producing data (Green and Thorogood 2004). The interviewer was a university-based researcher with no involvement in FP assessment or the Trust employing participants. This impartiality enabled her to approach data collection from a neutral standpoint without having to bracket any preconceptions about what would be found in the study. Participants were made aware of this which was felt to encourage openness and trust between the two parties. There were no personal characteristics of the researcher that were considered to influence the interaction with interviewees or the data obtained.

Interviews were held in confidential locations on the hospital sites with just the researcher and interviewee present and lasted between 30 and 50 minutes. The study was approved by the local research ethics committee. Consent

forms were signed at the start of the interview and all participants agreed to the audio taping of the interview.

### Data analysis

Interviews were transcribed by a member of the university research support team and checked for accuracy by the researcher prior to data analysis. The data analysis was led by the researcher supported by two colleagues on the research team: one a senior academic colleague with a shared role in postgraduate training for doctors (SG), and one a senior member of the FP management team at the Deanery (KF). Thematic content analysis as described by Green and Thorogood (2004) was used to guide the process of analysis, an approach widely used in health research to categorise and describe significant themes arising in an interview. Consistent with this approach, an initial coding scheme was devised based on the interview topics and reading early transcripts, and then all data were coded by the researcher with the support of the qualitative analysis package QSR Nvivo (Gibbs 2002). This generated five top level codes each having between three and nine sub-codes. Following this, re-reading and checking of coded transcripts was carried out, including testing out alternative coding. Coding was reviewed for face validity by the two further members of the research team (SG and KF). To improve rigour, one transcript was re-coded by the researcher four weeks later to check for consistency in the application of codes (Green and Thorogood 2004) with a high level of consistency found.

The final stage was the compilation of coded data into main themes which was

carried out by the researcher. Coded extracts were sorted and compared to build categories of data that conveyed similar meaning, taking into account distinct perspectives of trainees and assessors. This resulted in four main themes with final interpretation of themes including referral back to original transcripts and consultation with the research team. Extraction and compilation of the coded excerpts under theme headings was also completed using NVivo. This aided the search for deviant cases which are reported in the Findings.

### Findings

A selection of findings relating to the more significant themes are discussed here and illustrated with selected quotes from participants.

#### *Benefits of the FP assessments*

Trainees and assessors both recognise the main potential benefit of FP assessment as the opportunity for giving and receiving feedback as reflected in this quote by an assessor:

*“The greatest potential benefit is lots of contact with senior doctors and junior doctors, not directly related to patient care but sitting down separately to talk about a doctor’s training...” (Assessor)*

Trainees particularly value the legitimate access to consultant time and the benefits this leads to in terms of informal learning and teaching opportunities, as well as confidence building:

*“..it empowers you to demand that you get something from your seniors...a means of getting some senior teaching, which has been the most rewarding part.” (Trainee)*

However, there was a lack of enthusiasm among the trainees for the system overall, and scepticism among assessors about the benefits being realised. Many comparisons were made to the 'old system' revealing the perception that little had changed. As one trainee explained:

*"I just feel like I am a House Officer, the same as people were last year... I may be F1 but it's just new language, I'm still doing the same job just with a big folder of paperwork..." (Trainee)*

All the assessors and some trainees remarked on the potential value of the assessments as a 'safety net' to catch a failing trainee early on. There were, however, doubts as to whether this was yet being achieved, and concerns about lack of guidance on what action to take with a trainee who showed signs of failing.

### **Workload and organisation**

Assessors and trainees were in agreement that the assessment system as a whole was straightforward:

*"On top of starting a new job, having this big assessment thing was quite daunting. But once, you know, you're a month or two into the job, you realise it's actually really not that much at all." (Trainee).*

Other comments about administration and the forms used also showed agreement among the interviewees that these issues presented no problems, with just one assessor having strong negative opinions of the 'bureaucracy' involved.

However, a major theme in discussions was time pressures. Trainees reported difficulties in organising assessments and

securing the time and commitment of assessors, which required considerable perseverance as this trainee explains:

*"I spent about four weeks chasing one consultant and he kept putting it off and off and off until I went 'you're doing it now! Come on!' and he ticked any old box." (Trainee)*

Assessors felt that the time required to complete assessments 'properly' was an added work pressure. They reported a lack of time spent with individual trainees, leading to insufficient knowledge of trainees on which to base reliable judgements. One assessor reflected:

*"I think doing it well is quite difficult because we don't see enough of the trainees and to do it really well would take quite a lot of time." (Assessor)*

There was a sense of frustration at having insufficient time for the process, and arising from this, evidence of a negative 'tick box' attitude as seen above and discussed below in context of rigour.

*"Most of the time the people I've spoken to I've said I've got to do another mini-CEX and they've said oh, just give me the form, tick, tick, tick, tick, tick and signed it at the bottom." (Trainee)*

Reasons cited by assessors for such negativity were lack of interest, experience, training or time.

### **Rigour**

#### ***Validity of tools***

The tools used in assessments are seen to capture the appropriate competencies for doctors at this stage of training. The most popular tool for trainees was the Mini-PAT, due to the open feedback

this generates and relevance to everyday working, with this quote serving as a powerful reminder of the link from assessment to daily practice and patient care:

*“..the mini-PAT is the all-round winner because it looks at things which are more important to our colleagues and patients in the way that we behave and interact, just doing our job I suppose.” (Trainee)*

Next in popularity was the CbD because of the opportunity for on-the-spot training by the assessor. However, there was some reporting of duplication in assessments. DOPS was perceived by most to duplicate skills tested in medical finals and therefore to be ‘pointless’ and ‘condescending’. The Mini-CEX was seen to have significant overlap with CbD and to be a considerable drain on assessors’ time.

### **Scoring system**

A key area of concern for both trainees and assessors was the rating scale used in all assessments to summarise the outcome (a range of scores from one to six, where one is ‘below expectations’ and six is ‘above expectations’). Trainees recognise that marks are rarely awarded across the full range of the scale, and feel that the system does not encourage people to excel:

*“..everyone comes up with very much the same results... I just feel it sort of encourages mediocrity, in a way..” (Trainee)*

*“..it offers very little to the brilliant person.... there’s no facility within it to identify someone whose performance is outstanding...” (Assessor)*

One assessor admitted to: “giving everyone slightly above average marks”, giving the reason that trainees ‘don’t like to hear’ they are just average.

This may be one factor which fuels the ‘tick box’ attitudes mentioned above, a consistent theme in comments by both trainees and assessors. This quote is a typical view:

*“Most of the time the people I’ve spoken to I’ve said I’ve got to do another mini-CEX and they’ve said oh, just give me the form, tick, tick, tick, tick, tick and signed it at the bottom.” (Trainee)*

The transparency of negative attitudes like this further undermines the rigour of the system and the valuing of assessment.

### **Subjectivity**

Both trainees and assessors were very concerned about the potential for bias from trainees being able to choose assessors for each assessment. One trainee referred to ‘getting your mates to do it’ and six others referred to ‘how well’ you get on with your assessor affecting the scores. The following comment reflects this attitude:

*“I think you’d be a fool if you took it to anyone who you knew didn’t like you or thought you weren’t that good..” (Trainee)*

Assessment was therefore seen to be highly subjective and related to relationships between staff.

### **Discussion**

The study captured rich data from a typical group of trainees and assessors in a large acute hospital in the UK involved in the first year of the new assessment

system in postgraduate training. There was very strong consistency in the themes arising from interviews and the findings are consistent with those from other UK studies (Grant et al 2005, Carr 2006) and with the recent MMC Inquiry (Tooke et al 2008), giving further credence to the data achieved.

There was clear consensus amongst trainees and assessors that the assessments have the potential to deliver some important benefits to newly qualified doctors. These include building relationships between junior and senior staff, generating structured feedback on performance, building confidence in doctors facing their first experiences as responsible clinicians, and providing a safety net to identify a potentially failing doctor at an early stage in time to take remedial action. These findings are consistent with those from an earlier evaluation of the pilot stage of FP assessment, which found the system was valued for its educational benefits and its role in delivering feedback to trainees (Grant et al 2005).

However, what this study reveals is that to realise these important benefits time barriers must be overcome. This was a conclusion also reached by Grant et al (2005) in the earlier evaluation but the data reported here shed light on the nature of these barriers. In particular, trainees struggle to secure the time of senior staff, especially consultants, and spend a good deal of time trying to organise assessments. Consultants feel they do not have enough time with trainees to get to know them prior to assessments, due to shorter attachments, also reported by Carr (2006). Assessors also feel they lack time to do assessments

properly, leading to hasty judgements based on impression.

In addition the study reveals some concern about the assessment tools, which have potential to give rise to a lack of confidence in the system. Although the tools used are generally accepted as having face validity, the judgements they generate are not. The ability for trainees to choose their own assessors is seen to result in highly subjective judgements; assessors rarely give low marks, and there is a tendency to give average scores, with no-one failing or excelling. Trainees are left with the perception of rapidly ticked boxes rather than meaningful feedback about competence and do not believe that excellence is acknowledged or encouraged, by the system. Tooke et al (2008) similarly found 'a 'tick box' perception prevails' among staff and considerable 'scepticism regarding the competency assessments employed' including the sense of replicating skills from medical school. One of Tooke et al's (2008) overarching conclusions was that the current system is 'unlikely to encourage or reward striving for excellence'

The present study has revealed a number of problem areas in competence assessment and the data inform a number of suggestions for improvement. These strategies may be useful to those planning the implementation of new postgraduate assessment systems, to avoid what Fox and Arnold (2008) term 'supervisor disengagement and burnout'. Firstly, the introduction of the system in the UK did not allow time for preparation of staff for new roles in assessment or the culture change in performance management, which

may in part explain some of the poor attitudes and practice revealed in the data. Preparing staff adequately for new roles in assessment could include the provision of mandatory training for assessors at Trust level, to improve understanding of assessment and ensure consistency with appropriate use of scoring and feedback systems, leading to early recognition of both high and low achieving trainees. Coupled with a training strategy, it would be essential for assessors to have protected time for assessment to ensure adequate time to attend training and to prepare for and complete assessments. Tooke et al (2008) similarly highlighted the need to grasp 'the training implications in terms of protected time, staff development and understanding of contemporary methods of assessment'.

Since this study the independent body responsible for postgraduate medical education and training in the UK, has produced clear standards for trainers which acknowledges the importance of time for the training role. (PMETB 2009a). These standards highlight the role of regulatory and professional bodies, as well as employing organisations, in making strategies such as training and protected time successful. A further recent statement in the UK issued jointly by all such relevant bodies, re-affirms a commitment to standards for trainers including the need for formal recognition of the training role for doctors in secondary care, which has not previously occurred (PMETB 2009b). How such developments will be audited is not yet known but the importance of quality assurance processes has been highlighted in specialist training in general practice where similar workplace-

assessment has recently been introduced. (Mamelok 2009).

Additional strategies to support training might include the introduction of a register of trained assessors to help trainees locate informed and, crucially, willing individuals from their workplace including those from nursing and other allied health professions, taking pressure off senior medical staff. Hays (2007), also seeking lessons from the UK changes, similarly recommends medical education be 'nurtured' as a career pathway, with individuals on that pathway needing specific preparation and support. In the UK, it has traditionally been only trainers in the primary care sector who have been given training and recognition for the role but the statement referred to above is evidence that this is set to change. (PMETB 2009b)

The data here also suggest that it is important to consider the optimum number of assessments used in any system and what each adds to the whole. The volume of work and with it the pressures on both trainees and supervisors could be managed by careful selection and piloting of tools which avoid overlap and duplication. This could minimise the risk of assessment fatigue seen in this study which erodes quality and with it, perceived validity of the system overall. A final strategy implicated in this data is to minimise the degree of bias which damaged the credibility of judgements for the trainees in this study. To achieve this, it would be appropriate to limit the degree of choice in allocating assessors, to ensure that some assessments are carried out by people who are not immediate co-



workers of the trainee. Even though some assessors must inevitably have sufficient familiarity with the trainee and their work it is not essential for all and the inclusion of some more distant colleagues would enhance objectivity and rigour.

## Conclusions

Competency assessment is an important element of postgraduate training with the potential to improve the performance of doctors. Its role in postgraduate medical education is expanding in the UK and elsewhere. This study found through in-depth methods, agreement that such assessment can potentially bring huge benefits in terms of improving feedback to trainees, building relationships between staff and providing a safety net for struggling trainees, which is especially important when the number of graduates is set to rise. However, problems with lack of time and training considerably undermine the educational value. The data carry lessons for Australia and New Zealand where similar modes of assessment are anticipated. The most significant is the importance of preparing staff for new roles in assessment and specifically, training of assessors including those from other health professions. This, combined with protected time for assessment and the role of assessor, will require the full commitment of professional and regulatory bodies in postgraduate training, education providers and employing organisations. Further strategies include optimising the number of assessments that are required, and maximising objectivity in the system by reducing the scope for trainees to select assessors. Such measures will help

to smooth the waters of change facing postgraduate training and minimise the burden on supervisors and trainees alike, ultimately improving educational outcomes.

## References

- Archer JC, Davies H (2004) On reflection, *Health Service Journal* 114, 5903: 26-27.
- Beard J, Strachan A, Davies H, Patterson F, Stark P, Ball S, Taylor P, Thomas S (2005) Developing an education and assessment framework for the Foundation Programme, *Medical Education* 39: 841-851.
- Carr S, (2006). The Foundation Programme assessment tools: an opportunity to enhance feedback to trainees?, *Journal of Postgraduate Medicine* 82: 576-579.
- Department of Health (2004) *Modernising Medical Careers – The Next Steps*. London, Department of Health.
- Fox, GJ and Arnold, SJ (2008) The rising tide of medical graduates: how will postgraduate training be affected?. *Medical Journal of Australia* 189: 515-518
- General Medical Council (2001) *Good Medical Practice*. London, General Medical Council.
- Gibbs G, (2002) *Qualitative Data Analysis – Explorations with NVivo*. Berkshire, Open University Press.
- Grant J, Maxted M, Chambers K, Owen H (2005) *Foundation Assessment Programme: qualitative evaluation for development of the assessment system*. Milton Keynes, Open University Centre for Education in Medicine.

- Gray S, Feest K, Sabey, A (2009) Letter to the Editor: Assessment in the Foundation Programme: a qualitative study of consultants and junior doctors in the South West of England. *Medical Teacher* 31, 1: 66.
- Green J, Thorogood N (2004) *Qualitative Methods for Health Research*. London, Sage.
- Hays, RB (2007) Reforming medical education in the United Kingdom: lessons for Australia and New Zealand. *Medical Journal of Australia* 187, 7: 400-403.
- Holmboe ES, Huot S, Chung J, Norcini J, Hawkins RE (2003) Construct validity of the MiniClinical Evaluation Exercise (MiniCEX). *Academic Medicine* 78, 8: 826-830.
- Holmboe ES, Yepes M, Williams F, Huot S (2004) Feedback and the Mini Clinical Evaluation Exercise. *Journal of General Internal Medicine* 19: 558-561.
- Mamelok, J (2009) Workplace-based assessment (WPBA) portfolios in licensing for general practice specialty training. *Education for Primary Care* vol. 20: 139-142.
- Norcini J, Blank L, Duffy D, Fortna GS (2003) The Mini-CEX: a method for assessing clinical skills. *Annals of Internal Medicine* 138: 476-481.
- Postgraduate Medical Education and Training Board (2009a) *Generic standards for training: version 1.1, September 2009*. London: PMETB
- Postgraduate Medical Education and Training Board (2009b) *Joint statement from PMETB, AORMC, GMC, COPMED and NHS Employers in relation to the role and recognition of supervisors of postgraduate medical education and training*. London: PMETB.
- Southgate L, Hays RB, Norcini J, Mulholland H, Ayers B, Woolliscroft J, Cusimano M, McAvoy P, Ainsworth M, Haist S, Campbell M (2001a) Setting performance standards for medical practice: a theoretical framework. *Medical Education* 35: 472-481.
- Southgate L, Cox J, David T, Hatch D, Howes A, Johnson, N, Jolly, B, Macdonald E, McAvoy P, McCrorie P, Turner J (2001b) The General Medical Council's Performance Procedures: peer review of performance in the workplace. *Medical Education* 35, (Suppl.1): 9-19.
- Tooke J, Ashtiany S, Carter D, Cole A, Michael J, Rashid A, Smith PC, Tomlinson S, Petty-Saphon K (2008) *Aspiring to Excellence: findings and final recommendations of the Independent Inquiry into Modernising Medical Careers*. London, Aldridge Press.
- Violato C, Lockyer J, Fidler H (2003) Multisource feedback: a method of assessing surgical practice. *British Medical Journal* 326: 546-548.
- Wilkinson J, Benjamin A, Wade W (2003) Assessing the performance of doctors in training. *British Medical Journal* 327: s91-92.