



## **Principles for an assessment system for postgraduate medical training**

*A working paper from the Postgraduate Medical Education Training Board.<sup>1</sup>*

**Lesley Southgate**

Chair Assessment Committee

**Janet Grant**

Chair Curriculum Sub-Committee

14 September 2004

---

<sup>1</sup> This paper follows on from, and replaces, Principles and Standards for an Assessment System for Postgraduate Medical Training, prepared before the establishment of the PMETB. It has been adopted by the Board as a working paper for the Assessment Committee.

## Introduction

In this paper, an assessment system refers to an integrated set of assessments which is in place for the entire postgraduate training programme and which supports the curriculum. It may comprise different methods, and be implemented either as national examinations, or as assessments in the workplace. The balance between these two approaches principally relates to the relationship between competence and performance. Competence (can do) is necessary but not sufficient for performance (does do), and as experience increases so performance based assessment in the workplace becomes more important.

### Principle 1

The assessment system must be fit for a range of purposes

- The overall purpose of the assessment system must be documented and in the public domain
- The purposes of each and all components of the assessment system must be specified and available to the trainees, educators, employers, professional bodies including the regulatory bodies, and the public
- The sequence of assessments must match the progression through the career pathway
- Individual assessments within the system should add unique information and build on previous assessments

*The examples of purposes for assessment listed below are presented in the order that a trainee might encounter them. Position in the list does not imply importance<sup>2</sup>, neither is the list comprehensive.*

1. to inform career selection and choice
2. to confirm suitability of choice at an early stage of chosen career path

---

<sup>2</sup>*Sometimes purposes for assessment are not made explicit. Two examples are included here which can cause difficulties if all parties are not informed of the intention to use outcomes for these purposes.*

*1. to monitor the effectiveness of the training programmes 2. to identify the best people for posts where there is competition*

3. to demonstrate readiness to progress to the next stage of training having met the required standard
4. to provide feedback to the trainee about progress and learning needs
5. to support trainees to progress at their own pace by measuring progress in achieving competencies for chosen career path
6. to identify trainees who should change direction or leave medicine
7. to enable the trainee to collect all necessary evidence for revalidation
8. to assure the public that the trainee is ready for unsupervised practice
9. to provide evidence for the award of a CCST
10. to drive learning
11. to gain membership or fellowship of a medical Royal College or specialist association/society

## **Principle 2**

The content of the assessment will be based on curricula for postgraduate training which themselves are referenced to all of the areas of *Good Medical Practice*

- The programme will be based on an overall specification of the content for all component assessments, in order to avoid overlap or gaps
- Assessments will together systematically sample the entire content, appropriate to the stage of training, with reference to the common and important clinical problems that the trainee will encounter in the workplace and to the wider base of knowledge, skills and attitudes that doctors require
- The blueprint<sup>3</sup> from which assessments in the workplace or national examinations are drawn will be available to trainees and educators in addition to assessors/examiners

---

<sup>3</sup> The blueprint for an assessment specifies the content from which the assessment sample is drawn. In postgraduate medical training it usually comprises a matrix with one dimension broadly based on *Good Medical Practice*, and the other on the clinical problems that a trainee should be able to manage, at that stage of training.

### **Principle 3**

The methods used within the programme will be selected in the light of the purpose and content of that component of the assessment framework.

- Methods will be chosen on the basis of validity, reliability, feasibility, cost effectiveness, opportunities for feedback, and impact on learning.
- The rationale for the choice of each assessment method will be documented and evidence-based

*Large scale competence tests (e.g., MRCP, MRCGP, MRCPsych)*

- Approaches to the development and piloting of test items/clinical skills assessments for national tests of competence will be documented and available for external quality assurance. Studies to establish the validity of new methods will be undertaken.
- Systematic data collection will support the routine reporting of the reliability of tests of competence in high stakes pass/fail examinations. These statistics will be in the public domain.

*Work-based assessments (e.g., direct observation of consulting, 360 deg assessment, and case based discussions)*

- Must be subject to reliability and validity measures
- Evidence must be collected and documented systematically
- Evidence must be judged against pre-determined published criteria
- The weight placed on different sources of evidence must be determined by the blueprint and the quality of the evidence
- The synthesis of the evidence and the process of judging it must be made explicit

## Methods for Workplace Based Assessment

*For example:*

Systematic observation of clinical practice

*Direct Observation*

*Video*

Judgements of multiple assessors

Consulting with simulated patients

Case Record Review: including OPD letters

Case based discussions

Oral presentations

360 deg peer assessment

Patient surveys

Audit Projects

Critical incident review

### **Principle 4**

The methods used to set standards for classification of trainee's performance/competence must be transparent and in the public domain

- Standards in tests of competence such as national Royal College examinations, will be set using recognised methods based on test content and the judgments of competent assessors
- Where the purpose of the test is to provide a pass/fail decision, information from the performance of reference groups of peers should inform, but not determine, the standard
- The precision of the pass/fail decision must be reported on the basis of data about the test. The purpose of the test must determine how the error around the pass/fail level affects decisions about borderline candidates.
- Reasons for choosing either pass/fail or rank ordering should be described.
- Standards for determining successful completion of training to CST level should be explicit

## **Principle 5**

### **Assessments must provide relevant feedback**

- The policy and process for providing feedback to trainees following assessments must be documented and in the public domain<sup>4</sup>
- The form of feedback must match the purpose of the assessment
- Outcomes from assessments must be used to provide feedback on the effectiveness of education and training where consent from all interested parties has been given

## **Principle 6**

### **Assessors/examiners will be recruited against criteria for performing the tasks they undertake**

- The roles of assessors/examiners will be specified and used as the basis for recruitment and appointment
- Assessors or examiners must demonstrate their ability to undertake the role
- Assessors/examiners should only assess in areas where they have competence.
- The relevant professional experience of assessors should be greater than that of candidates being assessed
- Equality and diversity training will be a core component of any assessor/examiner training programme

---

<sup>4</sup> Sometimes it is appropriate to provide no feedback other than the test result. If this is a policy decision then reasons should be stated.

## **Principle 7**

There will be Lay input in the development of assessment

- Lay opinion will be sought in relation to appropriate aspects of the development, implementation and use of assessments for classification of candidates.
- Lay people may act as assessors/examiners for areas of competence they are capable of assessing.

## **Principle 8**

Documentation will be standardised and accessible nationally.

- Documentation will record the results and consequences of assessments and the trainee's progress through the assessment system
- Information will be recorded in a form that allows disclosure and appropriate access, within the confines of data protection.
- Uniform documentation will be suitable not only for recording progress through the assessment system but also for submission for purposes of registration and performance review
- Documentation should provide evidence for revalidation and compliance with *Good Medical Practice*
- Documentation should be transferable and accessible as the trainee moves location
- Documentation should be comprehensive and accessible to both to the trainee and to those responsible for training.

## **Principle 9**

There will be resources sufficient to support assessment

- Resources will be made available for the proper training of assessors
- Resources and expertise will be made available to develop and implement appropriate assessment methods
- Resources will support the assessment of trainees at national and local levels
- Appropriate infrastructure at national, deanery and Trust levels will support assessment.

## References

*This list of references is intended to provide a resource for considering the evidence which underpins this paper. It is not comprehensive.*

### Journal articles

1. Boulet JR, McKinley DW, Norcini JJ, Whelan GP Assessing the comparability of standardized patient and physician evaluations of clinical skills. *Adv Health Sci Educ Theory Pract.* 2002;7(2):85-97.
2. Cottrell D, Kilminster S, Jolly B, Grant J. What is effective supervision and how does it happen? A critical incident study. *Med Educ.* 2002 Nov;36(11):1042-9
3. De Grave WS, Dolmans DH, Van Der Vleuten CP. Student perspectives on critical incidents in the tutorial group. *Adv Health Sci Educ Theory Pract.* 2002;7(3):201-9
4. Fowell SL, Southgate LJ, Bligh JG Evaluating assessment: the missing link? *Med Educ.* 1999 Apr;33(4):276-81
5. Gorter S, Rethans JJ, van der Heijde D, Scherpbier A, Houben H, van der Vleuten C, van der Linden S. Reproducibility of clinical performance assessment in practice using incognito standardized patients. *Med Educ.* 2002 Sep;36(9):827-32
6. Gorter S, van der Heijde DM, van der Linden S, Houben H, Rethans JJ, Scherpbier AJ, van der Vleuten CP. Psoriatic arthritis: performance of rheumatologists in daily practice. *Ann Rheum Dis.* 2002 Mar;61(3):219-24.
7. Gorter SL, Rethans JJ, Scherpbier AJ, van Der Linden S, van Santen-Hoeufft MH, van Der Heijde DM, Houben HH, van Der Vleuten CP How to introduce incognito standardized patients into outpatient clinics of specialists in rheumatology. *Med Teach.* 2001 Mar;23(2):138-144
8. Grant J, Kilminster S, Jolly B, Cottrell D Clinical supervision of SpRs: where does it happen, when does it happen and is it effective? *Specialist registrars.* *Med Educ.* 2003 Feb;37(2):140-8.
9. Hays RB, Jolly BC, Caldon LJ, McCrorie P, McAvoy PA, McManus IC, Rethans JJ. Is insight important? measuring capacity to change performance. *Med Educ.* 2002 Oct;36(10):965-71
10. Heard SR, Roberts C, Furrows SJ, Kelsey M, Southgate L Corridor consultations and the medical microbiological record: is patient safety at risk? *J Clin Pathol.* 2003 Jan;56(1):43-7.
11. Hutchinson L, Aitken P, Hayes T. Are medical postgraduate certification processes valid? A systematic review of the published evidence. *Med Educ.* 2002 Jan;36(1):73-91.
12. Jolly BC, Ayers B, Macdonald MM, Armstrong P, Chalmers AH, Roberts G, Southgate LJ The reproducibility of assessing radiological reporting: studies from the development of the General Medical Council's Performance Procedures. *Med Educ.* 2001 Dec;35 Suppl 1:36-44
13. Kramer A, Muijtjens A, Jansen K, Dusman H, Tan L, van der Vleuten C Comparison of a rational and an empirical standard setting procedure for an OSCE. Objective structured clinical examinations. *Med Educ.* 2003 Feb;37(2):132-9.
14. Kramer AW, Dusman H, Tan LH, Jansen KJ, Grol RP, van der Vleuten CP. Effect of extension of postgraduate training in general practice on the acquisition of knowledge of trainees. *Fam Pract.* 2003 Apr;20(2):207-12

15. McAvoy PA, McCrorie P, Jolly B, Ayers AB, Cox J, Howes AD, Macdonald EB, Slimmon DJ, Southgate L Training the assessors for the General Medical Council's Performance Procedures.  
Med Educ. 2001 Dec;35 Suppl 1:29-35
16. Melnick DE, Asch DA, Blackmore DE, Klass DJ, Norcini JJ. Conceptual challenges in tailoring physician performance assessment to individual practice.  
Med Educ. 2002 Oct;36(10):931-5
17. Norcini J. The validity of long cases.  
Med Educ. 2001 Aug;35(8):720-1
18. Norcini JJ Setting XX on educational tests.  
Med Educ. 2003 May;37(5):464-9.
19. Norcini JJ Work based assessment.  
BMJ. 2003 Apr 5;326(7392):753-5.
20. Norcini JJ, Blank LL, Arnold GK, Kimball HR. Examiner Differences in the Mini-Cex.  
Adv Health Sci Educ Theory Pract. 1996;2(1):27-33
21. Norcini JJ, Blank LL, Duffy FD, Fortna GS The mini-CEX: a method for assessing clinical skills.  
Ann Intern Med. 2003 Mar 18;138(6):476-81
22. Norcini JJ, Lipner RS, Kimball HR Certifying examination performance and patient outcomes following acute myocardial infarction.  
Med Educ. 2002 Sep;36(9):853-9.
23. Prescott LE, Norcini JJ, McKinlay P, Rennie JS Facing the challenges of competency-based assessment of postgraduate dental training: Longitudinal Evaluation of Performance (LEP).  
Med Educ. 2002 Jan;36(1):92-7.
24. Rethans JJ, Norcini JJ, Baron-Maldonado M, Blackmore D, Jolly BC, LaDuca T, Lew S, Page GG, Southgate LJ. The relationship between competence and performance: implications for assessing practice performance.  
Med Educ. 2002 Oct;36(10):901-9
25. Roberts C, Sarangi S, Southgate L, Wakeford R, Wass V. Oral examinations-equal opportunities, ethnicity, and fairness in the MRCGP.  
BMJ. 2000 Feb 5;320(7231):370-5.
26. Schuwirth LW, Southgate L, Page GG, Paget NS, Lescop JM, Lew SR, Wade WB, Baron-Maldonado M When enough is enough: a conceptual basis for fair and defensible practice performance assessment.  
Med Educ. 2002 Oct;36(10):925-30
27. Schuwirth LW, van der Vleuten CP. ABC of learning and teaching in medicine: Written assessment.  
BMJ. 2003 Mar 22;326(7390):643-5.
28. SGorter S, Rethans JJ, Scherpbier A, van der Heijde D, Houben H, van der Vleuten C, van der Linden S. Developing case-specific checklists for standardized-patient-based assessments in internal medicine: A review of the literature.  
Acad Med. 2000 Nov;75(11):1130-7.
29. Southgate L, Campbell M, Cox J, Foulkes J, Jolly B, McCrorie P, Tomblinson P The General Medical Council's Performance Procedures: the development and implementation of tests of competence with examples from general practice
30. Southgate L, Cox J, David T, Hatch D, Howes A, Johnson N, Jolly B, Macdonald E, McAvoy P, McCrorie P, Turner J. The General Medical Council's Performance Procedures: peer review of performance in the workplace.  
Med Educ. 2001 Dec; 35 Suppl 1:9-19.
31. Southgate L, Hays RB, Norcini J, Mulholland H, Ayers B, Woolliscroft J, Cusimano M, McAvoy P, Ainsworth M, Haist S, Campbell M Setting performance XX for medical

- practice: a theoretical framework.  
Med Educ. 2001 May;35(5):474-81
32. Swanson DB, Clauser BE, Case SM. Clinical Skills Assessment with Standardized Patients in High-Stakes Tests: A Framework for Thinking about Score Precision, Equating, and Security.  
Adv Health Sci Educ Theory Pract. 1999;4(1):67-106
  33. SWragg A, Wade W, Fuller G, Cowan G, Mills P Assessing the performance of specialist registrars.  
Clin Med. 2003 Mar-Apr;3(2):131-4
  34. Tamblyn R, Abrahamowicz M, Dauphinee WD, Hanley JA, Norcini J, Girard N, Grand'Maison P, Brailovsky C. Association between licensure examination scores and practice in primary care.  
JAMA. 2002 Dec 18;288(23):3019-26.
  35. Van Dalen J, Prince CJ, Scherpbier AJ, Van Der Vleuten CP Evaluating Communication Skills.  
Adv Health Sci Educ Theory Pract. 1998;3(3):187-195
  36. Wass V, Wakeford R, Neighbour R, Van der Vleuten C Achieving acceptable reliability in oral examinations: an analysis of the Royal College of General Practitioners membership examination's oral component.  
Med Educ. 2003 Feb;37(2):126-31

## Books

### **International Handbook of Research in Medical Education**

edited by

**Geoffrey R. Norman**

*McMaster University, Hamilton, Canada*

**Cees P.M. van der Vleuten**

*Maastricht University, The Netherlands*

**David I. Newble**

*University of Sheffield, UK*

Kluwer Academic Publishers, Dordrecht

Hardbound, ISBN 1-4020-0466-4

Jolly, B and Grant, Janet [1997] **The Good Assessment Guide**. Open University Centre for Education in Medicine. ISBN 1 873207 76 X

### **Thanks to**

*John Norcini, Dave Swanson, Brian Jolly, Chris McManus, Geoff Norman, David Newble, Lambert Schuwirth for their helpful comments*