

# Designing Assessment Programmes for the Model Curriculum for Emergency Medicine Specialists

On behalf of the Core Curriculum and Education Committee for the International Federation for Emergency Medicine:

James Kwan, MBBS\*<sup>†</sup>; Nicholas Jouriles, MD<sup>†</sup>; Andrew Singer, MBBS<sup>‡</sup>; Venkataraman Anantharaman, MBBS (S'pore)<sup>§</sup>; Gautam Bodiwala, CBE, DL, DSc (Hon), MS<sup>¶</sup>; Peter Cameron, MBBS\*\*<sup>¶</sup>; C. James (Jim) Holliman, MD<sup>††</sup>; Cees van der Vleuten, PhD<sup>††</sup>; Cherri Hobgood, MD<sup>§§</sup>

**Keywords:** graduate medical education, curriculum, assessment, IFEM

## INTRODUCTION

There has been a proliferation of competency-based postgraduate training programmes in emergency medicine (EM) worldwide, including Australia, Canada, Singapore, the United Kingdom, and the United States. Several competency frameworks have been developed at national and international levels as a basis for competency-based postgraduate training programmes. These frameworks include the Accreditation Council of Graduate Medical Education (ACGME) competencies in the United States,<sup>1</sup> the Canadian Medical Educational Directives for Specialists (CanMEDS),<sup>2</sup> and Common Competences for Emergency Medicine in the United Kingdom<sup>3,4</sup>. In response to this increased emphasis on competency-based education during postgraduate training in EM, the International Federation for Emergency Medicine (IFEM) has recently developed a model curriculum to define the basic minimum standards for specialist training in EM.<sup>5</sup> The goal for specialist training in EM is to ensure that its trainees develop the necessary knowledge, skills, and professional attitudes to provide safe, expert, and independent emergency care within their own country. Accurate assessment of trainees' progress during specialist

training is of paramount importance to the educational process.

In recent years there have been changes to the assessment process. Traditionally, assessment has been considered exclusively as a process of measuring whether trainees have acquired the necessary knowledge, skills, and professional attitudes to practice independently as a specialist in EM, or assessment of learning. However, it is now recognized that an equally important function of assessment is to stimulate the individual's learning process—in other words, assessment *for* learning.<sup>6</sup> This new paradigm of the role of assessment should be firmly embedded in the educational process.

As a result of this conceptual change in assessment, there has been a shift from considering individual methods to programmes of assessment, to allow adequate sampling of performance of complex competencies in authentic contexts.<sup>7</sup>

## DESIGNING PROGRAMMES OF ASSESSMENT IN EM

A theoretical model for the design of a programme of assessment for competency-based curricula in postgraduate medical education has been described.<sup>8</sup> This model provides a practical approach to designing assessment programmes based on 10 principles of best practice. The General Medical Council in the United Kingdom has integrated these principles into a

From the \*University of Western Sydney, Sydney, Australia; †Akron General Medical Center, Akron, Ohio, USA; ‡Australian Government Department of Health and Ageing, Canberra, Australia; §Singapore General Hospital, Singapore, Singapore; ¶International Federation for Emergency Medicine; ¶De Montfort University, Leicester, UK; \*\*The Alfred Hospital Emergency and Trauma Centre, Monash University, Melbourne, Australia; ††Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA; ††Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands; and §§Indiana University School of Medicine, Indianapolis, Indiana, USA.

**Correspondence to:** Cherri Hobgood, Faculty Office Building, 720 Eskenazi Avenue, Indianapolis, IN 46202; Email: chobgood@iu.edu

quality management checklist to evaluate the quality of workplace-based assessment programmes in post-graduate training.<sup>9</sup> These principles will be used to develop an assessment framework for the IFEM model curriculum for EM specialists. An outline of this assessment framework is provided in Appendix 1.

## **PROPOSAL FOR AN IFEM ASSESSMENT FRAMEWORK FOR SPECIALIST TRAINING**

### **1. Define the purpose of assessment**

The purpose of assessment must be clear and transparent to the trainees being assessed. Ideally, the purpose of assessment is two-fold: (i) assessment of learning to determine if the trainee has attained the necessary competencies outlined in the model curriculum for independent practice as an EM specialist at the completion of a training programme; (ii) assessment for learning to encourage learning with the provision of feedback to improve the trainee's performance during training. The IFEM recognizes that this may not initially be possible in countries with limited EM infrastructure and educational resources.

### **2. Select an overarching competency framework**

Prior to selecting an overarching competency framework, a working definition of EM competence is required. Although several definitions for competence exist in the literature, these definitions are difficult to apply in actual clinical practice. At the simplest level, the Oxford Dictionary of English defines competence or competency as the "ability to do something successfully." In applying this definition to EM, competence of an EM specialist: (i) is based on the ability to perform specific clinical tasks; (ii) requires the integration of both domain-dependent (medical knowledge) and domain-independent (communication skills and professionalism) competencies in the domain of EM; (iii) is measurable in terms of observable behavior; (iv) is specific to the context of EM.<sup>10,11</sup>

A useful framework to consider in relation to learning outcomes is Miller's pyramid.<sup>12</sup> The development of competence requires the progression from the knowledge components of competence ("knows"), to the integration of behavioural elements of competence in a simulated environment ("shows how"), and into

the actual workplace ("does"). The application of this pyramid to the model curriculum requires assessment to test the performance of EM trainees in the workplace. In other words, what the trainee actually "does" in the workplace.<sup>13</sup> Although the goal of assessment should be focused at the summit of the pyramid to ensure the validity of the assessment programme, the lower levels of the pyramid ("knows," "knows how," "shows how") should not be ignored.

Furthermore, assessment must address both domain-dependent and domain-independent competencies. Several major educational frameworks for competency-based curricula have emerged to address the assessment of this additional dimension.<sup>3,14,15</sup> The IFEM does not endorse a specific framework and recommends that member nations evaluate the utility of these frameworks within their own countries.

### **3. Define the progression from novice to expert**

The assessment programme should acknowledge the progress of a trainee during training, from novice through to the development of competency and expertise. There has been a lack of understanding of how the competencies needed to perform complex tasks develop over time.<sup>16</sup> The ACGME, in collaboration with the American Board of Emergency Medicine, has embarked on a milestone project to identify the behaviours and attributes that describe the competencies with performance standards at the completion of each year of residency training.<sup>17</sup> These milestones provide a developmental roadmap for competencies and sub-competencies during specialist training, with the aim of ensuring that the trainee is being assessed at the appropriate level and informing trainees of their progress during training.

The IFEM recommends that IFEM nations identify these developmental milestones within their own programme of assessment.

### **4. Design a blueprint of the curriculum**

The test content must be carefully mapped to the learning outcomes. This is known as "blueprinting." The design of this blueprint should reflect the educational goals of the curriculum and ensure adequate sampling of the curriculum content if fair and reliable assessments are to be achieved.

The blueprinting process requires: (i) identification of a conceptual framework, such as the ACGME, CanMEDS and Good Medical Practice frameworks, to provide a structure against which to plan the content of the assessment programme; and (ii) careful mapping of the content within the framework to ensure broad sampling of the curriculum. An example of this blueprinting process is presented in the online supplementary appendices (see Appendices 2-7). Importantly, the milestones of development during training identified earlier should be incorporated into the blueprint to ensure that trainees are assessed at the appropriate level. The curriculum blueprint also provides an overview of the curriculum to ensure that appropriate assessment tools are chosen.

The IFEM does not endorse a specific conceptual framework for this mapping process, and IFEM nations may choose to create a new conceptual framework or adopt existing frameworks for the blueprinting process. Irrespective of this choice, there should be broad sampling of the curriculum to ensure that students are comprehensively and fairly assessed.

### **5. Select appropriate assessment methods**

The selection of assessment tools requires careful consideration of the utility of individual assessment methods within a programme of assessment. A conceptual model which defines the utility of an assessment method in terms of its validity, reliability or reproducibility, acceptability, feasibility, and impact on learning (educational effect) can be used to guide the selection of assessment methods.<sup>7</sup> Furthermore, additional criteria have been identified in the context of competency-based assessment programmes.<sup>18-20</sup> The IFEM acknowledges that the weight that a specific criterion carries within a programme of assessment will vary according to the purpose of assessment and each criterion's importance as perceived by those responsible for assessment within IFEM nations. Despite the importance of a high level of validity and reliability of a particular assessment method, both acceptability to stakeholders and feasibility in terms of available resources are also likely to be weighted heavily. The educational effect is one that has great importance when considering the role of assessment for learning. Therefore, the choice of one assessment method over another will in most instances be a compromise, depending on the specific assessment context and purpose.

The assessment of what the trainee actually “does” in the workplace is the goal of assessment for the IFEM model curriculum. This must be borne in mind when considering the choice of assessment methods. In addition, the assessment methods best suited to measure different competencies should be constructively aligned with the curriculum.<sup>21</sup> An example of assessment methods that are currently recommended for use by the ACGME<sup>22</sup> and the CanMEDS<sup>23</sup> frameworks, mapped to competencies and learning outcomes defined by the IFEM model curriculum, is available online (see Appendices 6 and 7). The IFEM does not advocate the superiority of one assessment method over another, but recommends that IFEM nations consider the utility of methods in the context of the educational resources that are available in their own country.

There are essentially three main categories of assessment methods, related to their position in Miller's pyramid:

a. Knowledge-based assessments (“knows,” “knows how”)

These methods consist of written, computer-assisted, and oral examinations. If appropriately designed, these methods have the ability to assess both factual knowledge and the application of knowledge (clinical reasoning, decision-making). Examples of written and computer-assisted examinations include both selected response formats, such as single best answer multiple choice questions, extended matching questions, and constructed response formats, such as short answer questions, essays, and script concordance tests.<sup>24</sup> Examples of oral examinations include the structured oral exam and chart-stimulated recall exam. The chart-stimulated recall examination has the potential to assess clinical reasoning and decision-making in the clinical environment, and therefore could potentially have the ability to assess performance in the workplace.

b. Performance-based assessment in a simulated environment (“shows how”)

These methods assess the performance of a trainee in a simulated environment. Examples of these methods include Objective Structured Clinical Examination (OSCE), Standardized Patients, and Simulation.

c. Performance-based assessment in the workplace (“does”)

These methods assess the performance of a trainee in the actual clinical environment through direct observation. Common examples include Mini Clinical Evaluation Exercise (Mini-CEX), Direct

Observation of Procedural Skills (DOPS), in-training evaluation reports, shift report cards, Multi-Source Feedback (MSF), patient surveys, logs, clinical record reviews, and portfolios.

Although it is beyond the scope of this paper to describe the utility of these methods in detail, a summary of the utility of some of the assessment methods described above is presented in the supplementary online appendices (Appendices 8-10). In addition, several excellent reviews currently exist which describe the utility of these assessment methods in general postgraduate training<sup>8,9,25-27</sup> and with reference to the ACGME<sup>21,22,28-33</sup> and CanMEDS framework<sup>10,23</sup>.

## **6. Decide on the stakes of the assessment**

Traditionally, there has been an artificial divide between formative and summative assessment. Formative assessment is typically low-stakes, and is intended to promote learning by providing feedback to trainees to identify their strengths and weaknesses in order to help improve their future performance and map their progress through training. Applying the criteria for good assessment,<sup>20</sup> it is clear that the validity, educational effect, and catalytic effect are of primary importance. Feasibility plays an important role, as a significant time commitment from both trainers and trainees is required to ensure that feedback is ongoing, timely, and tailored specifically for each trainee. Acceptability for both trainer and trainees is also important if they are to commit to the process and give credibility to the feedback provided.

In contrast, summative assessment is considered high-stakes, and is primarily intended to make end point decisions on whether a trainee is competent to independently practice as a specialist in EM. The emphasis is on validity, reproducibility, and equivalence, to ensure the credibility of the scores. Although acceptability, feasibility, and educational effect still play an important role, they do so to a lesser degree. Opportunities for feedback to improve future performance (catalytic effect) in this setting should not be neglected, and attempts should be made to integrate feedback into the assessment.

In a programme of assessment, the line between summative and formative assessment is blurred. There is a planned arrangement of individual assessments in a training programme. Expert judgment is imperative for

assessing complex competencies, and when there is a need to combine diverse information. Each assessment now represents a single data point, which is low-stakes, with an emphasis on stimulating learning through the provision of information-rich quantitative and qualitative feedback on the trainees' performance.<sup>34</sup> Individual data points can represent any assessment method at any level of Miller's pyramid. Higher-stakes assessments are based on the aggregation of many data points, with the number of data points being proportionate to the stakes of the assessment. Although quality compromises are made for individual methods, there is no compromise in the overall quality of the assessment programme. As a result, programmatic assessment will fit its purpose through integration with the learning programme and allow robust decision-making over a trainee's performance.

## **7. Involve stakeholders in the design of the assessment programme**

The IFEM recommends that all stakeholders involved in the educational process, including trainees, trainers, educational institutions, patients, health care systems, and regulators, should be actively involved in the design, standard setting, and evaluation of the assessment programme. It is important in the design of any assessment programme that accurate inferences can be drawn from a trainee's performance through well-defined and transparent processes.

Although a detailed consideration of standard setting is beyond the scope of this paper, two main processes for setting standards exist: norm and criterion referencing.<sup>35,36</sup> A balance of norm and criterion referencing should be used to set standards, which will be determined by the resources available and the consequences of pass/fail decisions in IFEM member nations.

Nevertheless, IFEM countries should ensure that the rationale for their standard-setting processes is credible, defensible, supported by a body of evidence, feasible, and acceptable to all stakeholders.<sup>37</sup>

## **8. Aggregation and triangulation of assessment results**

The measurement of a trainee's performance is a complex process. It requires the aggregation of assessment information from direct observation of the trainee's performance in the workplace—in other words what the trainee “does”—and from high-stakes,

competency-based assessment at the “knows how” and “shows how” level, such as multiple choice exam or key feature assessment and OSCE. Triangulation of information from these assessment options will avoid overreliance on summative examinations alone, and ensure that a firm, defensible decision about a trainee’s progress can be made. This links the formative value of assessment with its summative functions.

In a programme of assessment, the aggregate information is held against a performance standard and a committee of examiners makes a high-stakes decision. The decision will be based upon many data points, consisting of both quantitative and qualitative information. The trustworthiness of the decision is optimized through the use of procedural measures such as credibility, transferability, dependability, and confirmability, inspired by qualitative research methodologies.<sup>11,34</sup> Within a programme of assessment, the potential for additional redundant assessments is reduced, preventing assessor fatigue.

The portfolio has the potential to aggregate assessment information from different sources in larger programmes of assessment at the “does” level.<sup>38</sup> It also plays a role in stimulating learning through its reflective component. However, the implementation of portfolios can be a complex and resource-intensive process.

### 9. Assessor selection and training

Selection and training of assessors is required for a range of tasks, including providing feedback during formative assessment and making consistent, independent, defensible judgments during high-stakes summative assessments. The establishment of a formal training programme and calibration processes for assessors will improve the quality of assessment decisions. Professional advice and support for assessors through peer mentoring programmes is important to help new assessors develop confidence in identifying poor performance. The IFEM recommends that assessor training be a major component of any assessment programme, to ensure that assessors are capable of providing feedback and making appropriate decisions during summative assessments.

### 10. Quality improvement

The assessment programme should have quality assurance procedures in place to ensure that there is

continuous monitoring of the assessment programme, that adjustments can be made, and that there is constructive alignment of the assessment with the learning process. Importantly, processes should be in place to ensure that any changes made to the assessment criteria are explicit and clearly understood by all stakeholders, including assessors and trainees. A regular review of the test material is essential to improve the quality of the assessment.

A potential strategy to help less experienced IFEM member countries to develop a programme of assessment for specialist training in EM may be the establishment of regional hubs of expertise in assessment by IFEM member countries, with established specialist training programmes in EM.

**Competing interests:** None declared.

### SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/cem.2015.39>

### REFERENCES

1. Accreditation Council for Graduate Medical Education. *Accreditation Council for Graduate Medical Education (ACGME) Core Competencies*. Available at: [www.acgme.org/acWebsite/home/home.asp](http://www.acgme.org/acWebsite/home/home.asp) (accessed 11 October 2011).
2. The Royal College of Physicians and Surgeons of Canada. *The CanMEDS Physician Competency Framework*. 2005. Available from: <http://www.royalcollege.ca/public/canmeds/framework> (accessed June 2, 2012).
3. General Medical Council. *Good Medical Practice*. London: General Medical Council; 2009. Available at: [http://www.gmc-uk.org/static/documents/content/Good\\_medical\\_practice\\_-\\_English\\_0914.pdf](http://www.gmc-uk.org/static/documents/content/Good_medical_practice_-_English_0914.pdf).
4. The College of Emergency Medicine. *Curriculum and Assessment Systems for Core Specialty Training ACCS CT1-3 & Higher Specialty Training ST4-6 Training Programmes*. Available at: [http://www.gmc-uk.org/Curriculum\\_for\\_Emergency\\_medicine\\_2010\\_\\_amended\\_2012\\_.pdf\\_61015364.pdf](http://www.gmc-uk.org/Curriculum_for_Emergency_medicine_2010__amended_2012_.pdf_61015364.pdf).
5. Hobgood C, Anantharaman V, Bandiera G, et al. International Federation for Emergency Medicine model curriculum for emergency medicine specialists. *CJEM* 2011;13(2):109-21.
6. Schuwirth LW, van der Vleuten CP. Programmatic assessment: From assessment of learning to assessment for learning. *Med Teach* 2011;33(6):478-85.
7. van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ* 2005;39(3):309-17.
8. Wass V, van der Vleuten C. Assessment in medical education and training 1. In: *Medical Education and Training: From Theory to Delivery*. Carter Y, Jackson N editors. Oxford: Oxford University Press; 2009, 105-28.

9. General Medical Council Workplace Based Assessment: A guide for implementation; 2010. Available at: [http://www.gmc-uk.org/Workplace\\_Based\\_Assessment\\_\\_A\\_guide\\_for\\_implementation\\_0410.pdf\\_48905168.pdf](http://www.gmc-uk.org/Workplace_Based_Assessment__A_guide_for_implementation_0410.pdf_48905168.pdf).
10. Sherbino J, Bandiera G, Frank JR. Assessing competence in emergency medicine trainees: an overview of effective methodologies. *CJEM* 2008;10(4):365-71.
11. van der Vleuten CP, Schuwirth LW, Scheele F, et al. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol* 2010;24(6):703-19.
12. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(9 Suppl):S63-7.
13. Cydulka RK, Emerman CL, Jouriles NJ. Evaluation of resident performance and intensive bedside teaching during direct observation. *Acad Emerg Med* 1996;3(4):345-51.
14. ACGME. Accreditation Council for Graduate Medical Education (ACGME) Core Competencies. 1999. Available at: <http://www.acgme.org/acWebsite/home/home.asp> (accessed October 11, 2011).
15. Frank JR. CanMEDS: better standards, better physicians, better care; 2005. Available at: <http://www.royalcollege.ca/public/resources/aboutcanmeds> (accessed October 11, 2011).
16. Carraccio C, Burke AE. Beyond competencies and milestones: adding meaning through context. *J Grad Med Educ* 2010;2(3):419-22.
17. ACGME and the American Board of Emergency Medicine. The Emergency Medicine Milestone Project; 2012. Available at: <http://acgme.org/acgmeweb/Portals/0/PDFs/Milestones/EmergencyMedicineMilestones.pdf>
18. Messick S. The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher* 1994;23(2):13-23.
19. Baartman LKJ, Bastiaens TJ, Kirschner PA, van der Vleuten CPM. The wheel of competency assessment: Presenting quality criteria for competency assessment programs. *Studies in Educational Evaluation* 2006;32(2):153-70.
20. Norcini J, Anderson B, Bollela V, et al. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach* 2011;33(3):206-14.
21. Jouriles N, Burdick W, Hobgood C. Clinical Assessment in Emergency Medicine. *Academic Emergency Medicine* 2002;9(11):1289-94.
22. ACGME and American Board of Medical Specialties (ABMS). *Toolbox of Assessment Methods*. Version 1.1; 2000.
23. Bandiera G, Sherbino J, Frank JR, eds. The CanMEDS Assessment Tools Handbook: An Introductory Guide to Assessment Methods for the CanMEDS Competencies, 1st ed. Ottawa: The Royal College of Physicians and Surgeons of Canada; 2006.
24. Lubarsky S, Charlin B, Cook DA, et al. Script concordance testing: a review of published validity evidence. *Med Educ* 2011;45(4):329-38.
25. Boursicot K, Etheridge L, Setna Z, et al. Performance in assessment: consensus statement and recommendations from the Ottawa conference. *Med Teach* 2011;33(5):370-83.
26. Kogan JR, Holmboe ES, Hauer KE. Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. *JAMA* 2009;302(12):1316-26.
27. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach* 2007;29(9):855-71.
28. Dyne PL, Strauss RW, Rinnert S. Systems-based practice: the sixth core competency. *Acad Emerg Med* 2002;9(11):1270-7.
29. Hayden SR, Dufel S, Shih R. Definitions and competencies for practice-based learning and improvement. *Acad Emerg Med* 2002;9(11):1242-8.
30. Hobgood CD, Riviello RJ, Jouriles N, et al. Assessment of communication and interpersonal skills competencies. *Acad Emerg Med* 2002;9(11):1257-69.
31. King RW, Schiavone F, Counselman FL, et al. Patient care competency in emergency medicine graduate medical education: results of a consensus group on patient care. *Acad Emerg Med* 2002;9(11):1227-35.
32. Larkin GL, Binder L, Houry D, et al. Defining and evaluating professionalism: a core competency for graduate emergency medicine education. *Acad Emerg Med* 2002;9(11):1249-56.
33. Wagner MJ, Thomas HA. Application of the medical knowledge general competency to emergency medicine. *Acad Emerg Med* 2002;9(11):1236-41.
34. van der Vleuten CP, Schuwirth LW, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach* 2012;34(3):205-14.
35. Cusimano MD. Standard setting in medical education. *Acad Med* 1996;71(10 Suppl):S112-20.
36. Norcini JJ. Setting standards on educational tests. *Med Educ* 2003;37(5):464-9.
37. Schuwirth L, van der Vleuten C. How to design a useful test: the principles of assessment. In: *Understanding Medical Education: Evidence, Theory and Practice* (ed. Swanwick T). Oxford: Wiley-Blackwell; 2010, 195-207.
38. Van Tartwijk J, Driessen EW. Portfolios for assessment and learning: AMEE Guide no. 45. *Med Teach* 2009;31(9):790-801.