Developing a "Frequent Look and Rapid Remediation" Assessment System for a New Medical School

Chris Ricketts, PhD, and John Bligh, MA, MD

Abstract

The Peninsula College of Medicine and Dentistry (PCMD) was among the five new medical schools approved in the United Kingdom in 2000. PCMD required a new curriculum and a new outcomes-oriented assessment system based on sound educational theory. The resulting system was designed to use multiple sampling and show increasing authenticity as students progress through the curriculum. A "frequent look and rapid remediation" structure was incorporated to allow faculty to evaluate student competence throughout the year and take immediate action when warranted. Using a recently published framework, the authors retrospectively evaluate the assessment system and share many of the decisions they and their colleagues had to make. They discuss how to support stakeholders in understanding and contributing to the development of an assessment system and how to meet some of the challenges they encountered. They consider indicators of success in terms of the performance of the assessment system as a whole and the ways the system contributes to educational research. They suggest that future research should focus on building greater flexibility into the system so that assessment decisions can be individualized to particular students.

The literature about assessing clinical competence includes many articles on the development, performance, advantages, and disadvantages of various assessment instruments. Epstein and Hundert,[1] for example, undertook an extensive review of assessment methods, which they found to be wanting in some respects, and Fowell et al[2] identified problems specific to methods used in the United Kingdom. Few researchers, however, have evaluated the overall performance of an assessment scheme. The assumption seems to be that if the individual pieces are fit for purpose, then the system works as a whole. One recent study[3] discussed the need to think about assessment instruments as part of an overall assessment system which itself is a major part of the curriculum. Some researchers and some regulatory bodies[4,5]—in the United States, the

Accreditation Council for Graduate Medical Education; and, in the United Kingdom, the Postgraduate Medical Education and Training Board (now merged with the General Medical Council [GMC]) and the Quality Assurance Agency for Higher Education— aim to describe frameworks or standards for assessment programs, but a limited number of articles[6,7] describe working programs.

Many curricula and assessment systems divide the study of medicine into disciplines with preclinical and clinical content. During a nine-year period (April 2001 to March 2010), we and our colleagues developed and evaluated an assessment system in a new medical school where disciplinary boundaries are deliberately blurred and initial medical training is viewed holistically. In this article, we describe the development approach and decisions, changes made in response to internal and external feedback, and some of the outcomes selected to measure student performance. We hope that by sharing our experiences, we may help other organizations develop and improve their own assessment systems.

History and Context

In 2000, the creation of five new UK medical schools was approved[8]; among the five was the Peninsula College of Medicine and Dentistry (PCMD), a joint venture between the Universities of Exeter and Plymouth and five local National Health Service (NHS) trusts.

The GMC provides guidance to UK medical schools in the form of "Tomorrow's Doctors,"[9] a set of regularly revised medical school standards. GMC policy requires each education program to consist of both "core" and "student-selected" components. Students' performance on both components must be assessed.

Medical training in the United Kingdom is largely based on a five-year program that most students enter directly from (high) school at age 18. The new PCMD curriculum was designed to reflect the five-year structure and meet GMC standards. The first students enrolled in September 2002 and graduated in July 2007.

Developing a New Curriculum and System of Assessment

The PCMD curriculum was designed according to the principles of the PRISMS approach, which calls for medical curricula to be practice based, relevant to communities, interprofessional, shorter in length, continuous in nature, and taught across multiple sites.[10] The resulting curriculum takes an integrated view of medicine; specific "subjects" or "disciplines," such as anatomy[11] or pharmacology, are not taught via formal, single-topic courses. Rather, learning is based around patient presentations, and structured problem-

Dr. Ricketts was director of assessment, Peninsula College of Medicine and Dentistry, Plymouth, United Kingdom, and a National Teaching Fellow at the time this article was written. He has retired.

Prof. Bligh was vice dean for education, Peninsula College of Medicine and Dentistry, Plymouth, United Kingdom, at the time this article was written. He is now dean of medical education, Division of Medical Education, School of Medicine, Cardiff University, Cardiff, United Kingdom.

Correspondence should be addressed to Prof. Bligh, Division of Medical Education, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; telephone: (+44) 2920-743014; fax: (+44) 2920-743299; e-mail: blighj@cardiff.ac.uk.

based learning groups act as the main teaching modality in the first two years. The driving force behind the curriculum is producing students who are optimally prepared for practice in the NHS and who meet GMC requirements as set out in "Tomorrow's Doctors."[9]

Curriculum development focused on three key phases: the curriculum for years 1 and 2 (phase 1), the curriculum for years 3 and 4 (phase 2), and the curriculum for year 5 (phase 3). It took place against an overall longitudinal blueprint for the curriculum. Much of this development occurred simultaneously. While rollout took place year on year, development was an iterative process.

The assessment system was developed in three phases alongside the curriculum. The PCMD assessment scheme needed to follow the same principles of integration and clinical relevance as the curriculum and to align with the PRISMS model: "Assessment will emphasize 'doing' rather than 'knowing,' and continuous methods of formative assessment, such as portfolios, will predominate."[10] Assessment approaches also had to provide good feedback to students on a continuous basis and develop in authenticity during the five years of the program. Because we desired to assess overall program outcomes rather than simply performance in independent "courses," we kept three final outcomes at the heart of the assessment program: (1) applied knowledge of life and human sciences, (2) clinical skills, and (3) personal and professional development.

Rather than regarding these outcomes as units of teaching, we took the unusual step of defining them as units of assessment. Each unit of assessment became a "module" with associated credits, which allowed us to comply with the usual modular structure of UK undergraduate degrees.[12] The assessment modules did not need to link directly to units of teaching, but relevant learning for each assessment module could take place in a variety of contexts and experiences, from small-group tutorial sessions to patient encounters.

Quality Assurance

Our two parent universities created an overarching joint approval and review

board (JARB) to manage the quality assurance and course approval process. JARB's interacting committees and processes ensure timely development and annual review of the infrastructure, curriculum, resources, and the assessment system. The GMC, through its Quality Assurance of Basic Medical Education (QABME) program[13] and regular annual visits, provided a regulatory overview of the program's quality before granting final approval in 2005.

The PCMD curriculum is delivered at a number of localities—including three district hospitals, two university campuses, and more than 200 primary care practices—across the counties of Devon and Cornwall in southwest England. Students normally move among sites during the five-year program. To ensure consistency and quality of learning experiences at all sites, the curriculum was developed centrally but drew on expertise from all localities; it is now managed centrally but delivered locally. To ensure quality and consistency of assessment across the sites, the assessment program was also devised centrally; like the curriculum, it is managed centrally but delivered locally within agreed quality boundaries. Assessors are trained locally. For quality assurance purposes, it was crucial that local teams not devise their own assessment programs.

Assessment System Design

Dijkstra et al[5] propose a general framework for designing assessment systems and provide six areas that assessment designers should consider: goals, program in action, support, documenting, improving, and accounting. Their study had not yet been published when we devised our assessment system, but it serves as a useful model for the retrospective evaluation of our design processes. We will try to link our evaluation to these general areas, but it should be noted that these are not sequential steps; the last area, "accounting," includes establishing an evidence base and should be one of the first steps in designing a system.

Assessment goals: Purposes and outcomes

The primary purposes (goals) of our assessment system are to

1. improve student learning by providing the student with regular, appropriate, and, most important, timely feedback;

2. evaluate student knowledge, skills, and attitudes;

3. provide a mark or grade that enables a student's performance to be established in relation to program learning outcomes; and

4. provide evidence on the performance of the curriculum.

The primary outcome is to certify PCMD graduates fit for practice in the NHS in accordance with GMC regulations.

Assessment principles

It is generally accepted[3,14,15] that assessment programs need to take account of validity, reliability, impact, and resources, although these may be expressed slightly differently by different authors (e.g., usability, affordability). To implement what we considered to be best practices in assessment in accordance with these needs, we adopted seven principles:

1. Policies and procedures should be informed by the best evidence and relevant educational theory.

2. Assessment should be authentic and relevant to the major curriculum outcomes of clinical competence and professional competence.

3. A "frequent look and rapid remediation" approach should underpin the program.

4. Appropriate standard-setting methods for all assessments should be used.

5. A mixture of continuous, cumulative, and end-point assessments should be included.

6. Multiple sampling should be incorporated.

7. The performance of assessment activities and assessors must be evaluated.

At the time we developed our seven principles, the precepts of the Quality Assurance Agency for Higher Education (which, along with the GMC, regulates UK medical schools) were largely related to the institutional-level governance of assessment. Nevertheless, we considered those precepts, which included specific references to reliability and validity, fairness, and clarity of criteria.[16] We did not explicitly include fairness (equity in relation to equality and diversity) and openness to students among our principles because they

were already enshrined in the values of the whole medical school and did not need to be reiterated for the assessment system. All of our assessments are analyzed to investigate equity, and some interesting findings have been published, such as evidence showing that students with specific learning disabilities, including dyslexia, are not disadvantaged by multiple-choice tests and that the commonly allowed accommodations enable them to perform up to their capability.[17]

Establishing the evidence base

Having established our goals and principles, we needed to choose an appropriate mix of instruments using the "best evidence and relevant educational theory" that aligned with the PRISMS model.[10] We scoured the literature for best assessment practices that fit with our curriculum structure and assessment goals.

This work resulted in a document entitled "Undergraduate Assessment Evidence Base," in which we provided the justification for the instruments and sampling scheme we selected. We made the document available to the external bodies monitoring the development of our curriculum.

Combining our goals with the evidence base led us to select tools for assessing our three modules. Knowledge of human and life sciences would be assessed in an applied context using patient scenarios; a picture of longitudinal growth would be built using a progress testing approach.[18] Clinical skills would be assessed by a mixture of individual procedural assessments and a modified (integrated) objective structured clinical examination (OSCE). Personal and professional development would be assessed by a mixture of portfolios and

professional judgments. Table 1 shows the structure of our "frequent look" assessment system, which emphasizes "doing rather than knowing" wherever possible: Faculty observe students doing clinical tasks and judge their behaviors. The system is largely based on continuous assessment; there are few high-stakes, end-of-year examinations. The system increases in authenticity as students progress through the program by basing a large proportion of assessments in the later years on real patient encounters.

Implementing the Assessment System

Implementing the assessment system involved numerous parallel processes. First, we devised regulations for approval by our parent universities; these included appeals processes using approved governance structures and became part of PCMD's

Table 1

Medical Student Assessment Events by Year of Curriculum and Theme, Peninsula College of Medicine and Dentistry, Academic Years 2006–2007 Through 2008–2009

| Year of curriculum | Assessment theme* | | | |
	Formative	Applied knowledge	Professionalism	Clinical
Year 1	• 1 progress test • 3 knowledge assessments • 1 reflective essay • 3 professionalism judgments	• 3 progress tests • 1 end-of-year knowledge test	• 9 judgments • 2 reflective portfolio analyses • 2 academic reviews	• 6 skills lab clinical competencies
Year 2	• 3 knowledge assessments • 2 professionalism judgments	• 4 progress tests	• 9 judgments • 2 reflective portfolio analyses • 2 academic reviews	• 8 skills lab clinical competencies • 5 patient-based clinical competencies • End-of-year integrated clinical skills examination
Year 3	• 12 patient-based clinical competencies	• 4 progress tests	• 14 judgments • 2 reflective portfolio analyses • 4 academic reviews	• 7 skills lab clinical competencies • 27 patient-based clinical reasoning presentations
Year 4		• 4 progress tests	• 14 judgments • 2 reflective portfolio analyses • 4 academic reviews	• 7 skills lab clinical competencies • 12 patient-based clinical competencies • 27 patient-based clinical reasoning presentations • End-of-year integrated clinical skills examination
Year 5	• 2 patient-based clinical encounters • 12 patient-based clinical presentations	• 4 progress tests	• 4 multisource judgments • 1 reflective portfolio analysis • 4 academic reviews	• 7 skills lab clinical competencies • 8 patient-based clinical encounters • 8 patient-based clinical presentations • 19 observed patient-based practical skills

* Formative assessments familiarize students with an assessment instrument or provide feedback. Progress tests consist of 125 single-best-answer questions using clinical vignettes. Professionalism judgments are made by academic or clinical staff on the basis of observed behaviors in a variety of contexts. Clinical competencies are normally single procedures, such as history taking, physical examination, or blood pressure measurement. In the clinical skills lab, these may use simulated patients, mannequins, or other simulation equipment. Those described as "patient based" use real patients. The integrated clinical skills examination in year 2 uses simulated patients, whereas the one in year 4 uses real patients.

academic regulations and codes of practice. As the system developed, we produced a technical manual detailing how the assessment process works (with information on all individual assessment instruments and standard-setting methods) and how results are aggregated to make decisions about student competence. The technical manual, a live document that is revised annually, is available to all students and staff through our managed electronic learning environment, and by request from the authors. All changes are approved in a multistage process involving specialist assessment groups, the medical program management committee, the college education committee, and, when necessary, the JARB.

Next, we identified, appointed, and trained relevant staff to take lead responsibility for the assessments described in Table 1, and we ensured that all assessors were trained. We put into place administrative processes that allowed assessment data to be collected, disseminated, and acted on. We also developed assessment blueprints, chose appropriate standard-setting methods, and developed quality assurance processes.

Further, we had to decide how closely the medical school would control the timing of assessments. PCMD's student-centered approach suggested that in some areas, such as clinical procedures, students might be able to decide when they were ready to be assessed.

Because our goal was to create an evidence-based assessment system, we also put into place staff and systems for evaluating our assessment instruments using appropriate psychometric approaches. We therefore appointed a psychometrician to work full-time on assessment analysis for the school. The outcomes of these analyses were fed back into the quality assurance processes of the school through structures such as the PCMD Education Committee and the JARB. They were also made available to external quality assurance agencies, such as the GMC through its QABME process. The evidence the analyses provided allowed us to make informed decisions regarding changes to assessment criteria, assessment instruments, standard-setting methods, and assessor training as our experience with the system matured.

These implementation issues had to be addressed in each of the three phases of development. We took the opportunity to learn from, and to modify, existing processes in each phase. However, we needed to ensure that the assessment contexts increased in authenticity at each stage of training in order to produce new graduates who could operate as junior doctors with real patients at the end of the program. So, whereas much assessment in years 1 and 2 of the curriculum was based on simulated patients or simulated instruments in a safe environment, assessment in year 5 was based on encounters with real patients in the clinical environment whenever possible.

Intervention and Remediation

Because our assessment system provides information at frequent points, faculty are able to review students' progress throughout the year and intervene when any assessment identifies poor performance. We therefore instituted a process of regular academic review that identifies struggling students and those who need additional support. As the review process developed, we came to use the phrase "frequent look and rapid remediation" to describe our assessment program. We wrote procedures for remediation and formed a core "remediation team" that works with students whose performance is of particular concern. We have also undertaken research on the effect of remediation on subsequent student performance.[19]

External Review

The developing assessment system was continually reviewed by PCMD, our parent universities, the GMC, and external examiners. We were open with all our assessment results in aggregate form, including any analyses that suggested that all was not perfect. We successfully engaged our external reviewers as part of the development process, and thus we were able to expand the expertise available to the school.

We benefited from the additional evidence that the external reviewers offered. For example, students who were in the early cycles of OSCE examinations and did not know the general content areas in advance claimed that they were disadvantaged because this information was "leaked" to students later in the cycle by those who had completed the assessment. External reviewers provided evidence that knowing the general content area in advance did not confer an advantage. We were able to confirm this using our own data after the examinations, and now we routinely include this analysis as part of our internal review after such assessments. However, we also decided to give an indication of general station content as preparatory information for all students.

Outcomes

The outcomes from the development of our assessment system are a set of successful assessment modules that perform to the required standard, the ability to identify students able to progress, and research publications contributing to the wider understanding of medical assessment.

As noted above, the system uses continuous assessment (see Table 1), and remedial action is possible at any of the many assessment points. Global assessments on progression are made at the end of every academic year. The system has identified students who are not ready or able to progress from one year to the next. In many of these cases, students were able to repeat the year of study and then make satisfactory progress. Other students left the program voluntarily or were asked to leave. The progression figures for 2006–2009, averaged over three cohorts, are shown in Table 2. As the table highlights, progression standards are slightly more stringent between years 2 and 3 and between years 4 and 5 than at other transition points. This is deliberate; both of these transitions reflect a large increase in contact with and responsibility for patients. The success of our graduates continues to be evaluated, and a recent report[20] showed that our graduates reported that they felt at least as well prepared for working in the NHS as graduates of other medical schools.

Challenges and Conclusions

The biggest challenge facing any assessment development team is that the system devised by a small number of staff will be tested to destruction by a large number of students. Students are prime stakeholders in any assessment system, and development teams can learn from their reactions to assessment instruments and processes. Responding to unexpected problems can seem like fighting fires, but this important process should be acknowledged more

Table 2

Medical Students Who Progressed Between Stages, Peninsula College of Medicine and Dentistry, Academic Years 2006–2007 Through 2008–2009

Transition	No. completing the year	No. progressing to the next year	% Progressing
Year 1 to 2	646	611	94.6
Year 2 to 3	653	581	89.0
Year 3 to 4	538	507	94.3
Year 4 to 5	562	497	88.3
Graduated	390	380	97.4
Overall	2,789	2,576	92.4

widely as part of improving an assessment program. For example, in response to student requests, we changed the basis of the knowledge test at the end of year 1. Students were concerned that the test, which was blueprinted against the knowledge required of a junior doctor, did not accurately indicate what they had learned in their first year because their exposure to clinical learning was not extensive. In addition to using the existing blueprint, therefore, this early knowledge test now assesses performance toward achieving the learning outcomes generated by the students themselves in their problem-based learning sessions.

Some assessment systems allow students flexibility to choose the timing of certain assessments. Clearly, this cannot occur when all students must sit the same assessment at the same time (e.g., our progress tests). However, for some clinical skills procedures we initially sought to encourage students to understand when they might be competent and to choose the time of their assessment, allowing repeat attempts. It rapidly became clear that this practice was not sustainable. The majority of students chose to leave the assessment as late as possible, creating a bottleneck and adding to staff workloads.

Review of the assessment program continues, and there are ambitions both to improve it and to use it to provide research evidence of value to the medical education community. Challenges for the future include reviewing the assessment system with a view to

simplification, using generalizability theory as a common framework to evaluate and gain a deeper understanding of the value of the assessments, and building greater flexibility into the existing system so that assessment decisions can be individualized to particular students.

References

1 Epstein RM, Hundert EM. Defining and assessing professional competence. JAMA. 2002;287:226–235.

2 Fowell SL, Maudsley G, Maguire P, Leinster SJ, Bligh J. Student assessment in undergraduate medical education in the United Kingdom, 1998. Med Educ. 2000; 34(suppl 1):1–49.

3 Van der Vleuten CP, Schuwirth LW. Assessing professional competence: From methods to programmes. Med Educ. 2005;39:309–317.

4 Lew SR, Page GG, Schuwirth LW, et al. Establishing defensible programmes for assessing performance. Med Educ. 2002;36:936–941.

5 Dijkstra J, Van der Vleuten CP, Schuwirth LW. A new framework for designing programmes of assessment. Adv Health Sci Educ Theory Pract. 2010;15:379–393.

6 Smith SR, Dollase RH, Boss JA. Assessing students' performances in a competency-based curriculum. Acad Med. 2003;78:97–107. http://journals.lww.com/academicmedicine/fulltext/2003/01000/assessing_students_performances_in_a.19.aspx. Accessed September 23, 2010.

7 Littlefield JH, Harrington JT, Anthracite NE, Garman RE. A description and four-year analysis of a clinical clerkship evaluation system. J Med Educ. 1981;56:334–340.

8 Bligh J. More medical students for England. Med Educ. 2001;35:712–713.

9 General Medical Council. Tomorrow's Doctors. http://www.gmc-uk.org/education/undergraduate/tomorrows_doctors.asp. Accessed September 7, 2010.

10 Bligh J, Prideaux D, Parsell G. PRISMS: New educational strategies for medical education. Med Educ. 2001;35:520–521.

11 McLachlan JC, Bligh J, Bradley P, Searle J. Teaching anatomy without cadavers. Med Educ. 2004;38:418–424.

12 The Quality Assurance Agency for Higher Education. Higher Education Credit Framework for England: Guidance on Academic Credit Arrangements in Higher Education in England. http://www.qaa.ac.uk/england/credit/creditframework.pdf. Accessed September 23, 2010.

13 General Medical Council. Overview of the QABME program. http://www.gmc-uk.org/education/undergraduate/qabme_programme.asp. Accessed September 7, 2010.

14 Harlen W. Criteria for evaluating systems for student assessment. Studies Educ Eval. 2007;33:15–28.

15 Knight P. A briefing on key concepts: Formative and summative, criterion & norm-referenced assessment. York, UK: Higher Education Academy; 2001. http://www.bioscience.heacademy.ac.uk/ftp/Resources/gc/Assess7.rtf. Accessed November 1, 2010.

16 The Quality Assurance Agency for Higher Education. Section 6: Assessment of students. In: Code of Practice for the Assurance of Academic Quality and Standards in Higher Education. 2nd ed. http://www.qaa.ac.uk/academicinfrastructure/codeofpractice. Accessed September 23, 2010.

17 Ricketts C, Brice J, Coombes L. Are multiple choice tests fair to medical students with specific learning disabilities? Adv Health Sci Educ Theory Pract. 2010;15:265–275.

18 Freeman A, van der Vleuten C, Nouns Z, Ricketts C. Progress testing internationally. Med Teach. 2010;32:451–455.

19 Lewis T. Tailored feedback for struggling undergraduate medical students. In: Association for the Study of Medical Education. Abstracts of the 2009 Annual Scientific Meeting, July 15–17, 2009. Edinburgh, UK: Association for the Study of Medical Education; 2009:85. http://www.asme.org.uk/images/ABSTRACTS_09.pdf. Accessed October 1, 2010.

20 Brennan N, Corrigan O. Longitudinal Evaluation of South West Peninsula Deanery F1 Doctors: Final Report. http://sites.pcmd.ac.uk/ice. Accessed September 2, 2010.