

# Learning in practice

---

## Using computers for assessment in medicine

Peter Cantillon, Bill Irish, David Sales

Computer based testing can offer many advantages over traditional paper based methods of assessment. The authors look at what it means and its potential uses for assessment in medicine

Department of  
General Practice,  
NUI Galway,  
Republic of Ireland  
Peter Cantillon  
*senior lecturer in  
general practice*

Examination  
Department, Royal  
College of General  
Practitioners,  
London SW7 1PU  
Bill Irish  
*MRCGP examiner*  
David Sales  
*deputy convenor*

Correspondence to:  
P Cantillon  
peter.cantillon@  
nuigalway.ie

BMJ 2004;329:606-9

Using computers for assessment in medicine is not a new idea, but, as information technology has become ever more important for teaching medicine, so computers have become an established means of student assessment.<sup>1</sup> Computer based testing (also called computer based assessment or computer assisted assessment) is not just an alternative method for delivering examinations, it represents an important qualitative shift away from traditional methods such as paper based tests. In this article we consider what is meant by computer based testing, its different manifestations, and its potential uses.

### What is computer based testing?

Computers are now regularly used to deliver, mark, and analyse student assessments.<sup>2</sup> There are two main types of computer based testing. The type familiar to many students is where candidates fill in their responses on a paper form, which is fed into a computer optical mark reader. This reads the form, scores the paper, and may even report on the test reliability. The second type of computer based testing is where computers provide an assessment interface for students: they input their answers and receive feedback via a computer.

### Where did computer based testing come from?

Computer based tests have been used since the 1960s to test knowledge and problem solving skills.<sup>3</sup> The earliest versions were text based and typically consisted of factual questions for which there were definite right and wrong answers. Such testing was attractive because it was possible to automate marking, and the students could receive instant feedback. Until recently computer based testing has mainly been used for formative assessment (such as giving students feedback on performance during a course). However, it is now increasingly being used for end point, summative examinations (such as the US medical licensing exam). Its acceptability as a means of assessment in high stakes examinations is now well established.<sup>4 5</sup>

### The rationale for computer based testing

Initial interest in computer based testing was driven by the time efficiencies that it offered compared with standard paper based tests in formative assessment.<sup>3</sup> As it began to be used for summative assessment, establishing whether computer based testing's performance was comparable to that of paper based assessments became important. Several studies have shown that computer based tests do perform as well,<sup>4 5</sup> and student surveys have shown that such tests were frequently more popular than traditional tests.<sup>6 7</sup>

Computer based testing was introduced to the US medical licensing exam in 1999, primarily to address concerns about the security of material with the paper test, which had a worldwide administration.<sup>8</sup> Computer based testing was found to offer several advantages, including (a) the delivery of high quality images that were required for some board certification exams such as in orthopaedics, (b) efficient data collection for statistical analysis, (c) opportunities for rapid feedback to candidates, (d) automated assembly of tests, and (e) the delivery of patient management problems to assess performance in evolving patient management simulations.<sup>8 9</sup> Box 1 lists the main advantages and disadvantages of computer based testing.

### How is computer based testing used?

Computer based testing can be used at several points in a course depending on the purpose of the assessment.<sup>10</sup>

*Diagnostic assessment*—Early in a course teachers can assess students' prior knowledge

*Self assessment*—During a course, students can assess their skills to identify their own learning needs

*Formative assessment*—Students receive computer generated feedback on their performance, and teachers can measure the effectiveness of their teaching

*Summative assessment*—Candidates have to pass the test to progress in a course or gain a particular qualification or accreditation.

A recent survey of UK medical schools and royal colleges revealed a growing interest in the use of computer based testing (unpublished data from electronic

survey by W Irish and P Cantillon, 2004). Most institutions already use it for formative student tests, but some are starting to use it in a limited manner for summative testing (see box 2 for examples of computer based testing). The commonest concerns relate to finding computer suites large enough to examine large classes and the issue of exam security.

### Question design

Early computer based tests were restricted to objective, text based questions and answers.<sup>11</sup> Now, however, software for creating a variety of online assessments is widely available. Assessors can, for example, incorporate images, drawings, and multimedia into questions, thus increasing the potential for testing beyond that offered by traditional paper based tests. The Tripartite Interactive Assessment Delivery System (TRIADS) resource at the University of Derby has several of the different question types that can be designed with modern software, and box 3 lists many possible question formats.

An example of the free software available to develop computer based tests is available at University of Leicester's CASTLE website ([www.le.ac.uk/castle/](http://www.le.ac.uk/castle/)). Several commercial software packages are also available (such as Questionmark) that allow teachers to create and operate computer based testing tests with minimal training. Virtual learning environments, which are becoming common in medical schools and educational institutions as a means of managing the curriculum, often include the ability to create computer based tests.

### Authentic assessment

Computer based testing is used increasingly to deliver "authentic" assessments, which are designed to simulate real world tasks and scenarios. Authenticity is important because there is good evidence that learners are better at storing and retrieving knowledge if it is learnt in a relevant context. It is also more appropriate to assess competence in a situation resembling the context in which the skill is to be used.<sup>12</sup> Patient management problems are often used in computer based testing to assess problem solving skills. Such problems typically begin with a patient's presenting complaint. Candidates are then asked to select appropriate items of history, examination, and investigation before making a diagnosis and outlining a management plan. However, test designers have found that, as the ability to solve problems is often context specific, a large number of computer based patient management

### Box 1: Pros and cons of computer based testing

#### Educational aspects

##### Advantages

Saves time for teachers by marking automatically and doing factor analysis  
Students can receive instant feedback on performance  
Teachers can easily track performance of individual students  
Testing can be delivered simultaneously at multiple sites  
Teachers receive feedback and evaluation data to inform their course designs  
Greater variety of media (such as video, graphics, etc) and test types can be used compared with paper based tests

##### Disadvantages

Generally favours objective type tests (open answer questions are often difficult to mark with automated marking systems)  
Question banks require expertise in designing test items and appropriate data storage  
May distinguish unfairly between students on the basis of ability to use computers rather than knowledge of the test subject  
Invigilation may be difficult in centres where students sit side by side to work on the same test simultaneously  
Plagiarism is possible if a test is done on line with student access to the internet and email  
Requires very reliable and secure computer test delivery systems

#### Administrative aspects

##### Advantages

Allows truly objective marking without human biases  
Marking is automated and rapid  
Tests can be assembled quickly from computer stored question banks

##### Disadvantages

Is potentially expensive and time consuming to set up  
Hardware and network failures mean that back up examinations must be prepared  
Staff who design or invigilate tests need to be trained for the purpose

problems are usually necessary to reliably test a candidate's problem solving ability.<sup>12</sup>

### Computer simulation

Patient simulators (such as the Harvey simulator developed at the University of Miami) are becoming a regular feature in many undergraduate and postgraduate courses. These are similar in concept to the simulators used to train airline pilots. They can combine mechanical, audiovisual, and data resources to create realistic clinical presentations. Learners can interact with the simulator, making judgments and errors without the fear of causing harm.<sup>13</sup> As with other forms of computer based testing, simulator based assessments can offer immediate feedback and correction of errors. Simulations are increasingly used for authentic summative and formative assessments of clinical competence and represent a qualitative shift away from traditional paper based assessment.<sup>14</sup>

### Security issues

Computer based tests can be difficult to invigilate.<sup>15</sup> As computer screens are usually upright, candidates may be able to see each other's screens. If an examination is online, candidates may be able to access internet resources to find answers to questions posed and to email each other. Medical schools and private companies have established dedicated computer based assessment centres to improve security. These centres often minimise opportunities for copying by placing opaque screens between candidates and use highly

### Box 2: Examples of computer based testing

- School of Medicine, University of Birmingham. MedWeb computer assisted assessment (<http://medweb.bham.ac.uk/caa/mcq/>)—Formative assessment for medical science students
- Rus Dewey's Psych Web. "Quiz Yourself" section ([www.psywww.com/selfquiz/](http://www.psywww.com/selfquiz/))—Formative assessment for psychology students
- United States Medical Licensing Examination ([www.usmle.org/](http://www.usmle.org/))—Summative assessment
- Driving Standards Agency. Theory component of UK driving test ([www.dsa.gov.uk/](http://www.dsa.gov.uk/))—Summative assessment

**Box 3: Range of questions that can be used in computer based testing (adapted from Seale 2002<sup>2</sup>)**

**Tick box type**

- Multiple choice questions
- Yes/no questions
- True/false questions
- Sequencing questions
- Multiple response questions
- Ranking questions

**Free text type**

- Case studies
- Essay questions
- Justification questions
- Short answer questions
- Assertion-reason questions
- Completion questions

**Questions using images**

- Image identification questions
- Drag and drop
- Graphical hotspot questions
- Labelling and building questions
- Build up image questions
- Drawing questions

**Other formats**

- Field simulation
- Sore finger questions
- Matching questions
- Text/numerical questions
- Completion questions
- Matrix questions

Several of these question types are described in more detail in the resources section of the Computer Assisted Assessment Centre website ([www.caacentre.ac.uk](http://www.caacentre.ac.uk))

effective data security and assessment invigilation regimes. Online examinations are carried in environments where access to the internet and email is controlled. Commercial computer based testing centres are often used to facilitate simultaneous multi-site tests, thus shortening the travel time between home and the nearest test centre.

Additional security is provided by randomising the order of questions delivered to different students, hence the chances of copying from adjacent candidates attempting the same questions are minimised. As learners become more adept at finding relevant answers to questions on the internet, so software to detect plagiarism in open question assessments has improved and is becoming more widely available for teachers.

**Item banks and adaptive testing**

One clear advantage of computer based testing is that questions can be tagged and stored in electronic question banks. These can be used by teachers to create new tests with relative ease, and they can be programmed to randomly select questions to form a new test. It is therefore possible to offer each student a different test while assessing the same course content. This has

many possible benefits, the most obvious being the enhanced security offered by an assessment in which each student is set a different set of questions. Another benefit is that students themselves can use question banks to generate random formative tests to assess their own performance and learning needs.

An exciting development in question bank technology is the use of adaptive testing. Traditionally, tests are delivered in the same format to all candidates regardless of their ability. Adaptive testing allows a test to be pitched at the level of expertise of each candidate, so that if a candidate is scoring well he or she will be offered increasingly difficult questions. This allows for a shorter testing time and the use of fewer questions.<sup>16</sup> Computer adaptive tests offer good reliability over short testing times for a wide range of different abilities. However, such tests are extremely difficult to design and are often criticised for not covering a sufficient breadth of key subjects.

**Computer based testing and the testing of higher order thinking**

One of the major concerns about computer based testing is whether it can assess higher intellectual skills (such as the ability to critically evaluate different sources of information).<sup>11</sup> This arises from the fact that it is difficult to programme computers to reliably assess free text (natural language) answers (such as essays). This explains the predominance of objective tests in computer based testing.

However, it is not the question type that dictates what level of intellectual skill is tested, rather it is the content of the question that determines what type of competence is assessed.<sup>12 17 18</sup> It is therefore possible to test higher intellectual functions through more sophis-

**Box 4: Questions to guide the introduction of computer based testing<sup>10</sup>**

1. Will computer based testing add to the existing assessment regime?
2. Is the purpose of the proposed computer based testing formative or summative?
3. Is there sufficient time to learn about computer based testing, write the questions, design the test, learn how to use the testing system, and deliver the test? (start small)
4. Is there local technical support available to deliver computer based testing?
5. Is the local network and internet access sufficiently robust to manage a computer based test?
6. Are the local hardware and network backup arrangements sufficiently robust to manage formative and particularly summative computer based testing?
7. Can students' answers in a computer based test be stored securely?
8. Are there sufficient computers to test large numbers of candidates simultaneously (or in sequence)?
9. Do all of the students have sufficient computer skills to perform adequately in a computer based test?
10. Will students have an opportunity to practise with computer based testing before facing a summative computer based test?
11. How will the test be scored and what kind of test analysis is required?
12. How will the test be invigilated?

## Summary points

Computers have been used for assessment in medicine since the 1960s: computer based assessment began as a formative assessment tool but increasingly is being used for summative and high stakes examinations

Computer based testing offers many advantages over traditional paper based tests including automated marking and student feedback, multimedia question types, and efficient test assembly

Computer based tests can be delivered anywhere via a secure computer network and are increasingly invigilated in dedicated, computer based, assessment centres at some distance from the test source

Using computers for high stakes assessment can be expensive and usually requires considerable attention to assessment design and test security

icated questionnaire design and the use of simulations. Reliable objective assessment of the responses to open questions is possible within particular educational contexts.<sup>19</sup> However this remains an area of development and there are as yet no widely available assessment systems that can reliably score free text answers in different contexts.

## Implementing computer based testing

Computer based testing offers many educational and practical advantages over paper based tests, but implementation can be challenging and lengthy. Although most doctors in training and undergraduate students have a high degree of computer literacy, it is not yet 100%. Establishing computer based testing for high stake examinations is also costly. One solution is to share costs and resources between different institutions. In the United States, for example, several medical accrediting institutions have devolved their testing role to the National Board of Medical Examiners, which designs, delivers, and reports the assessments on behalf of the accrediting bodies. An example is the US medical licensing exam, much of which is delivered online to test centres all over the United States.

In Britain the establishment of the Postgraduate Medical Education and Training Board in 2003 may offer an opportunity for the rapid roll out of computer based testing in professional examinations. This will require an unprecedented degree of collaboration between examining bodies such as the royal colleges, but the experience of the National Board of Medical Examiners in the United States has shown

how worth while such collaboration and sharing of resources can be.

## Conclusion

Computer based assessment is an emerging technology, with great potential for improving the assessment of doctors and other health professionals. In addition to many practical advantages, computer based testing can facilitate the development of more valid assessments. Experience worldwide has shown it to be generally popular with candidates and efficient for marking and delivery. However, the costs and expertise necessary to use this technology should not be underestimated. The establishment of organisations such as the US National Board of Medical Examiners and perhaps the UK Postgraduate Medical Education and Training Board may offer a model for effective collaboration in the design and delivery of computer based tests.

Contributors and sources: PC wrote the article, prepared the survey of UK medical schools, and coordinated the literature review. BI conceived of the idea for the study with DS, contributed to writing the paper, and carried out a survey with PC of computer based training in UK medical schools. DS conceived of the idea for the paper with BI and contributed to both the literature review and the writing of the article.

Funding: None.

Competing interests: None declared.

- 1 Appel J, Freidman E, Fazio S, Kimmel J, Whelan A. Educational assessment guidelines: a clerkship directors in internal medicine commentary. *Am J Med* 2002;113:172-9.
- 2 Seale J. Using CAA to support student learning; LTSN Generic Centre 2002. [www.ltsn.ac.uk/application.asp?app=resources.asp&process=full\\_record&section=generic&tid=38](http://www.ltsn.ac.uk/application.asp?app=resources.asp&process=full_record&section=generic&tid=38) (accessed 18 Jul 2004).
- 3 Swets JD, Feurzeig W. Computer aided instruction. *Science* 1965;150:572.
- 4 Russell M, Haney W. Testing writing on computers: an experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives* 1997;5(3).
- 5 Wolfson PJ, Velosky JJ, Robeson MR, Maxwell KS. Administration of open-ended test questions by computer in a clerkship final examination. *Acad Med* 2001;76:835-9.
- 6 Ogilvie RW, Trusk TC, Blue AV. Students' attitudes towards computer testing in a basic science course. *Med Educ* 1999;33:828-31.
- 7 Bocij P, Greasley P. Can computer-based testing achieve quality and efficiency in assessment? *Int J Educ Technol* 1999;1:1-17.
- 8 Wei H. Computer based testing and the USMLE. *Medical Computing Today* 1999. [www.medicalcomputingtoday.com/0ausmle.html](http://www.medicalcomputingtoday.com/0ausmle.html) (accessed 4 Mar 2004).
- 9 Dillon GF, Clyman SG, Clauser BE, Margolis MJ. The introduction of computer-based case simulations into the USMLE. *Acad Med* 2002;77:S94-6.
- 10 Smart C. Computer assisted assessment guide. <http://toomolfileserv.bangor.ac.uk/elearning/guides/caa/caa-pdf> (accessed 18 Nov 2003).
- 11 Thelwall M. Computer-based assessment: a versatile educational tool. *Computers and Education* 2000;34:37-49.
- 12 Schuwirth LWT, Van Der Vleuten CPM. The use of simulations in assessment. *Med Educ* 2003;37(suppl 1):65-71.
- 13 Hegarty MK, Bloch MB. The use of simulators in intensive care training. *Curr Anaesth Crit Care* 2002;13:194-2000.
- 14 Guagnano MT, Merlitti D, Manigrasso MR, Pace-Palitti V, Sensi S. New medical licensing examination using computer-based case simulations and standardized patients. *Acad Med* 2002;77:87-90.
- 15 Whittington D. Technical issues and security. In: Brown S, Bull J, Race P, eds. *Computer assisted assessment in higher education*. London: SEDA, Kogan Page, 1999:21-7.
- 16 Kreiter CD, Ferguson K, Gruppen LD. Evaluating the usefulness of computerized adaptive testing for medical in-course assessment. *Acad Med* 1999;74:1125-8.
- 17 Ward WC. A comparison of free-response and multiple-choice forms of verbal aptitude tests. *Applied Psychological Measurement* 1982;6:1-11.
- 18 Norman GR, Smith EK, Powles AC, Rooney PJ, Henry NL, Dodd PE. Factors underlying performance on written tests of knowledge. *Med Educ* 1987;21:297-304.
- 19 Foubister SP, Michaelson GJ, Tomes N. Automatic assessment of elementary standard ML program using Ceilidh. *Journal of Computer Assisted Learning* 1997;13:99-108.

## Recommended reading

Bull J, McKenna C. *Blueprint for computer-assisted assessment*. London: Routledge Falmer, 2003