Medical schools, clinical colleges and other groups are committed to improving the measurement of trainees’ clinical skills by using specific assessments such as OSCEs (objective structured clinical examinations),1,2 simulated patients,1 mini-short cases3 or portfolios1,2 (the latter a collection of evidence of ability, such as supervisor reports, audit of procedures or publications). However, as Miller has noted,4 “no single assessment method can provide all the data required for judgement of anything so complex as the delivery of professional services by a successful physician”. Most of us contribute by assessing trainees as they work with us — so-called “in-training assessment”2,5,6 Our judgments are based on observing their performance (how they are “doing” the job) — ie, the highest level of Miller’s four-level clinical assessment pyramid4,5 (see “Tips 6”).7

What we need to judge is broad — covering clinical competence, communication and professional skills. Unless we plan in advance, we could find ourselves lost at the end of a 12-week attachment, not really sure how well trainees are doing in these areas. Although there are many problems with the reliability of in-training assessments,5,9 they are extensively used and there are strategies for improving their reliability.

How do we measure performance?5

There are several ways to measure performance:
• Outcomes — eg, patient outcomes. However, this is difficult, as many factors influence patient outcomes.
• Process — eg, how well trainees have carried out a task, communicated, assessed a patient or written in the notes.
• Volume — eg, how many procedures the trainee has done.

In most circumstances, we measure performance based on how well trainees are working (ie, the “process”, as noted above), which is feasible and simple. Measuring patient outcomes or volume of work is more difficult.

Challenges with in-training assessment6,8

• As raters, we aren’t very good at being objective. Comparing results across examiners shows we tend to be either “hawks” (marking hard) or “doves” (marking easily).
• We tend not to distinguish between items — if trainees perform well in one area, we tend to assess them well in other areas (the “halo” effect).
• Personality traits (eg, extroversion, introversion) or poor command of English may have either a positive or negative impact on our assessment, irrespective of the trainee’s ability.
• If we do the trainee assessment long after the actual training period has taken place, we tend to mark towards the mean.
• Interaction with the trainee is important. If you are both the teacher and assessor, marks tend to be higher.

How can we improve?8,10

• Be familiar with the outcomes expected for trainees — in clinical competence, communication and professionalism.
• Turn these outcomes into observable behaviours:
  • Clinical competence — observe trainees doing an examination or taking a history, test their knowledge, review the inpatient notes or discharge summaries;
  • Communication — observe trainees speaking to patients, and require them to present to you;
  • Professional skills — note punctuality, time-management skills, whether trainees can cope with responsibility and whether they are interested in learning.
• Set expectations at the beginning of the rotation. Get trainees to take some responsibility for the assessment, such as bringing case notes for discussion.
• Find “assessable moments”, such as on rounds, in which trainees examine or talk to the patients and you watch. Write down your thoughts at the time and accumulate results across the term.9
• Assess multiple events during the training period, to make assessment more reliable.2,9
• Involve multiple people — ask other doctors, nurses or patients for their opinions (“360° assessment”).8,10

Feedback

Perhaps more important than the assessment per se is using the information we have gathered to give feedback (such as in appraisal). In assessment, although rating by means of a global score (“overall pass”, “borderline” or “fail”) works well,9 junior medical officers also want detailed feedback, not simply broad comments like “overall, you are very good”.

Self-assessment

It is useful to encourage a habit of self-assessment.11 Children tend to overestimate their abilities, whereas adults underestimate their own abilities. Poor students often overestimate their abilities.
However, if feedback is given, a side effect is that we get better at our self-assessment. So, before giving your feedback, ask trainees to fill in the assessment form before you do (self-assessment), or ask how they feel they are going.

Acknowledgements
We would like to thank the teachers and participants in Teaching on the Run courses for their input, and the Medical Training Review Panel, Australian Department of Health and Ageing, for funding support.

Competing interests
None identified.

References

(Received 25 Feb 2005, accepted 27 Apr 2005)