Clinical skills assessment

M H KELLY

L M CAMPBELL

T S MURRAY

SUMMARY

Background. *From September 1996, all GP registrars completing vocational training in the United Kingdom must demonstrate competence by means of a four-part assessment procedure.*

Aim. *To look at the accuracy of one of the components of vocational training: the trainer's report.*

Method. *Seventy-five registrars completing their general practice training at the end of July 1997 were invited to take part in a practical skills workshop. Eight stations were designed to test practical skills and diagnostic interpretations that were included in the trainer's report, and a clinical vignette accompanied each task. The marking schedule used was developed from the minimum standards required in the trainer's report. Twenty-nine registrars (38%) took part in the workshop.*

Results. *Only one registrar passed all eight stations. The maximum number of stations failed by any one individual was five and this doctor was the only one of the sample to ultimately fail summative assessment. The majority of registrars failed by being unable to interpret clinical findings. Twenty-five registrars (86%) responded to the follow-up questionnaire. Of these, only six felt that the stations were unrealistic. All but two registrars had spent at least six months in their hospital training doing obstetrics and gynaecology but, in spite of this, only 31% of registrars were above minimum competence for vaginal and speculum examination.*

Conclusion. *With one exception, registrars passed all aspects of the trainer's report. Discrepancy was found between the trainer's report and the doctor's ability to carry out clinical procedures. There is an assumption that many of these clinical skills are being taught and assessed at undergraduate level and during the hospital component, but this cannot be taken for granted. Doubt must also be cast on whether the trainers are using the trainer's report appropriately, and whether this is a valid and reliable tool to identify skills deficient in registrars for summative assessment.*

Keywords: clinical skills; competence; summative assessment; trainer's report.

Introduction

THE General Medical Council has emphasized the importance of the development of clinical skills,[1] and many medical schools have responded by developing facilities for the development and assessment of these skills using realistic simulations.[2-4] It has been shown that these developments are necessary because of proven concerns regarding the clinical skills of junior hospital doctors.[5]

From September 1996, all GP registrars completing vocational training in the United Kingdom (UK) have been expected to demonstrate competence by means of a four component assessment procedure. The four components are the submission of a videotape of consultations, a multiple choice paper containing true/false and extended matching items, the submission of an audit project, and a trainer's report. The trainer's report is the only instrument that attempts to assess practical skills. This report has been described in detail elsewhere by Johnson *et al*.[6,7] However, this report allows the trainer to use indirect methods of assessing practical skills; e.g. by a trainer discussing the skills with the registrar.

Evidence is available that suggests that general practitioners (GPs) do not necessarily have adequate practical skills,[8] and it has been shown that self-assessment has a low correlation with performance-based assessment.[9] On the other hand, it has been shown that clinical skills are best assessed by direct observation.[10] We therefore set out to answer two questions:

1. Can GP registrars successfully carry out a series of common clinical procedures?
2. How does their performance in a skills workshop compare with their trainer's assessment of the same skills in the structured trainer's report?

Method

The 75 registrars completing their general practice training at the end of July 1997 were invited to attend a practical skills workshop. The workshop coincided with the completion of the summative assessment trainer's report.

Description of stations

Eight stations were set up to test practical skills and diagnostic interpretations included in the trainer's report. Standard medical simulations ('dummies') were used for vaginal and rectal examination, ophthalmology, ear examination, and injection techniques. At these four stations, clinical examination and interpretation of findings were required. The remaining stations were 'task-based' stations where the equipment used was well recognized as being reliable and valid for carrying out that specific task — peak flow meter; needles and syringes for injection; and an electronic blood pressure cuff and dummy arm with pre-set blood pressure measurement allowing distinction between the fourth and fifth heart sounds, which is part of the trainer's report. A clinical vignette accompanied each task (Table 1). Each station was set up by MK.

Marking schedule

The marking schedule used was developed from the minimum standards required in the trainer's report for each of the clinical skills and tasks.

Assessors

The assessors were experienced GPs who were either trainers or current MRCGP examiners. One observer was appointed to each

M H Kelly, FRCGP, assistant director of Postgraduate General Practice Education; and T S Murray, FRCGP, director of Postgraduate General Practice Education Department of Postgraduate Medical Education, Glasgow. L M Campbell, FRCGP, general practitioner, The Surgery, Kirkintilloch, Glasgow.

Submitted: 4 March 1998; final acceptance: 22 February 1999.

Table 1. Examples of clinical stations.

Station One	Clinical Scenario	
Using the peak flow meter	A 36-year-old female, who has been a well-controlled asthmatic, comes to your asthma clinic for a review. Estimate her peak flow and interpret the results.	
Assessment by Observer		
Doctor uses the correct equipment	Yes	No
Doctor teaches the patient how to use the meter correctly	Yes	No
Doctor uses and is able to use appropriate charts to interpret the results correctly	Yes	No
Doctor is able to apply the results to clinical practice	Yes	No
Has the registrar reached the standard for independent general practice?	Yes	No
Station Four	**Clinical Scenario**	
Vaginal examination and use of vaginal speculum	A 38-year-old lady attends with a history of heavy painful periods and several episodes of postcoital bleeding. Carry out the vaginal examination indicated and record your findings.	
Assessment by Observer		
Doctor is able to undertake bi-manual examination	Yes	No
Doctor is able to describe findings systematically	Yes	No
Doctor is able to interpret the findings (including ability to detect major abnormality)	Yes	No
Doctor uses clean speculum of appropriate size and uses gloves	Yes	No
Doctor inserts and removes speculum comfortably	Yes	No
Doctor visualizes the cervix	Yes	No
Doctor interprets the findings made	Yes	No
Has the registrar reached the standard for independent general practice?	Yes	No

station to ensure observer consistency, and each was individually briefed on the assessment expected. All candidates were given the same guidance and were reminded at each station to behave as if they were dealing with a patient in the real clinical setting. Each assessor was taken to each of the stations individually by MK and was asked to carry out the appropriate examination without the aid of a clinical scenario. At the four stations where clinical examination and interpretations of findings were required, each of the assessors, observed by MK, correctly identified the clinical findings.

A questionnaire was sent to the registrars after the workshop to ascertain their views on how useful and realistic the exercise had been for them. A further questionnaire was also sent to the trainers asking them to grade, for each of the clinical scenarios, their degree of confidence in their registrar being able to attain minimal competence.

Results

Twenty-nine (38%) registrars took part in the workshop: 14 male, 15 female. Only one registrar passed all eight stations. The maximum number of stations failed by any one individual was five: this doctor was the only one of our sample to fail summative assessment and failed both the video component and the trainer's report of summative assessment.

The performance at each station and the number of stations failed are shown in Tables 2 and 3. Stations 4, 5, and 8 had the highest number of failures. At station 4 (gynaecology), 17 registrars failed to identify a bulky uterine fibroid and five did not correctly identify a cervical polyp. Three registrars did not wear gloves to do a vaginal examination, one did not do a vaginal examination at all, and one failed to carry out a speculum examination.

At station 5 (anaphylaxis), the most common faults were failure to give adrenaline and failure to check the expiry date of the drug. Similar faults were found at station eight with regard to the expiry date, but three registrars also gave the wrong dose of tetanus. Of those who failed only one station, six failed station 4 (gynaecology), three failed station 5 (tetanus), and one failed station 8 (anaphylaxis).

Registrar feedback

Twenty-five registrars (86%) responded to the follow up questionnaire. Only six registrars thought that the stations were unrealistic and, of these, four mentioned the diabetic fundus, although all passed this station. All but two responders had spent at least six months of their hospital training doing obstetrics and gynaecology and had received certificates of competence (Table 4). In spite of this, only 31% of registrars were above minimal competence for a vaginal and speculum examination. All but two of the registrars found the workshop to be of value and felt it should be carried out earlier in the year.

Trainer confidence

Twenty-five of the 29 (86.2%) trainers returned the questionnaire. Twenty-three trainers felt confident, very confident, or extremely confident that their registrar would perform above minimal competence for each station. Of the remaining two trainers, one had little confidence that his registrar would adequately carry out a bi-manual examination and interpret the findings, and this was the case. The other trainer expressed doubt

Table 2. Number of stations failed.

Number of stations failed	Number of registrars (n = 29)
0	1
1	10
2	10
3	4
4	3
5	1
>5	0

Table 3. Number of registrars passing individual stations.

Station number		Number of registrars passing (%) (n = 29)
1	Using the peak flow meter	25 (86)
2	Taking a blood pressure	28 (97)
3	Use of ophthalmoscope	20 (67)
4	Vaginal examination and use of speculum	9 (31)
5	Giving an intravenous injection (anaphylaxis)	17 (58)
6	Use of an auroscope	29 (100)
7	Rectal examination	29 (100)
8	Giving an intramuscular injection (tetanus)	16 (55)

Table 4. Hospital experience of registrars.

Post (six months except where indicated)	Number of Registrars (n = 25)
Accident and emergency[b]	22
Dermatology	5
Infectious diseases	2
Geriatrics[c]	20
Medicine[a]	14
Obstetrics and gynaecology[d]	23
Orthopaedics	2
Paediatrics	8
Psychiatry	20

[a]One registrar spent two years in specialty, four spent one year; [b]two spent one year in specialty; [c]one spent 18 months in specialty; [d]spent 2 1/2 years in specialty.

that his registrar would be able to use an ophthalmoscope and interpret the findings. The registrar failed this station, along with four others, but the trainer did not express any doubt in any of these areas in the trainer's report.

Summative assessment trainer's report

With one exception, the registrars passed all aspects of the trainer's report. The registrar who failed summative assessment failed five of the eight stations but was deemed by the trainer to be above minimal competence in all of these areas in the report, with consultation and organizational skills being highlighted as areas of concern. Of the 20 who failed the station on vaginal and speculum examination, 18 had been assessed in the practice by direct observation, one by tutorial and one by discussion. Of the doctors who failed the injection and anaphylaxis station — two failed to specify how they had been assessed — two had been assessed by discussion and the remainder by direct observation.

Discussion

Our results show that a significant number of GP registrars were unable to correctly carry out a number of straightforward procedures despite receiving a satisfactory trainer's report. It has been shown[10] that physical examination skills are best assessed with precise measurement, which avoids the 'halo effect' of overall evaluation that occurs with subjective rating scales, and our results were obtained using direct observation. While it is relatively easy in practice to check examination technique, the report is less than ideal in situations where clinical abnormalities are present and does not allow for interpretation of findings to be judged. Instead it encourages trainers to rely on more nebulous means of assessment such as discussion with the registrar or the erroneous assumption that completion of a relevant hospital job confirms the presence of these skills. It could be argued that the results of our study were biased by the fact that not all eligible GP registrars took part in the study and that we did not use real patients but simulations. Our assessors were satisfied that the simulations were realistic. They undertook the stations where any clinical judgement was required, tested the reliability of the electronic blood pressure machine, and found consistent results. They were observed by one individual and were blinded to the clinical problem. They also observed the same station throughout the day to maximize reliability and minimize any variation. Clinical scenarios were only used for the candidates so that each station would be put in a clinical context and reduce criticism of artificiality. Only six registrars felt the stations were unrealistic and the criticisms were not related to station 4, which had the most failures.

Despite this, it may be that some doctors did not behave as they normally would despite being given clear instructions to do so. We must hope that is true in the example of the three doctors who did not wear gloves for a vaginal examination. Even allowing for behaviour differences, it cannot explain why 17 doctors failed to identify a large uterine fibroid that was obvious to all of the assessors.

Given the fact that only one registrar out of 29 was able to carry out the whole series of assessments successfully, it would suggest that the number of unsuccessful registrars in the whole population is likely to be substantial. As previously published work regarding summative assessment was mirrored throughout the UK, this suggests a major problem in the GP registrars nationwide.

In a recent publication on piloting the trainer's report, Johnson *et al*[11] admit that the trainer's report is less than ideal in assessing clinical skills, particularly those concerning intimate examinations. Their study also demonstrated an urgent need to clarify the guidance notes and for the layout to be revised. In fact, in an amended version they advocate the use of mannequins as a means of assessing clinical skills. More concerning was the finding that, of the eight registrars referred, five would have been missed in the current system. Given this evidence, it is our conclusion that the trainer's report should have been more rigorously evaluated before being included as a 'stand alone' component of summative assessment — a process that is designed to protect the public from the non-competent doctor.

We would propose that a larger study throughout the UK be undertaken to quantify the extent of the problem. Many of these tasks are legitimately seen by the trainers as being taught and assessed at undergraduate level; however, from our results it appears that such skills should not be taken for granted in doctors entering their vocational training year.

This study has emphasized the importance of direct observation in assessment of clinical skills and has shown the value of using mannequins. Although resource-intensive in terms of time and labour, the assessments were acceptable and realistic to the candidates and could be used formatively as a teaching tool as

well as summatively in assessment. For more intimate examinations these may be more practical to use than directly observing real patients. In particular, GP trainers should only complete the relevant sections of the trainer's report when they have evidence of competent performance. In addition, the attitude of hospital consultants to the issuing of form Vocational Training Regulation 2 has to be questioned. This form is a declaration that the registrar in question is competent in the appropriate aspects of the specialty. It would be interesting to know if consultants are aware of the significance of signing this document and what forms of assessment, if any, they use to inform their judgement.

This paper would suggest that hospital colleagues do little in the way of appraisal and allow the senior house officers (SHOs) to move on to the next post without being aware of their deficiencies and level of clinical competence. Large parts of the trainer's report could best be completed as part of formative assessment during the hospital posts, with the SHOs carrying their assessment with them to their registrar year.

It would be interesting to know how the competencies discussed were assessed at undergraduate level; for example, is the ability to carry out and interpret a vaginal examination assessed during undergraduate gynaecology?

Conclusion

Attention should be paid to the ability of medical students and junior hospital doctors to carry out procedures. There is a need for a more rigorous assessment structure in the hospital sector, with some form of record of performance or log book that could be carried forward by junior doctors demonstrating that their performance of procedures had been observed and approved. GP trainers should make no assumptions about the skills of their registrars and therefore cannot abdicate their responsibility to test these skills either by direct observation in the general practice context or by using a series of simulations. In particular, the ability of the registrars to detect common and important abnormalities should be assessed.

Summative assessment has been developed to overcome the weaknesses of a system that relied on the subjective opinion of a trainer and to give a more objective evaluation of registrars' competence. This paper has highlighted the flawed system of assessment that exists at undergraduate level and in the hospital sector. It is therefore vital that, in the practice year, summative assessment has a robust tool to assess clinical skills. It would appear that the trainer's report is unreliable at identifying skills-deficient registrars, and this component should be revisited.

References

1. General Medical Council. *Tomorrow's doctors: recommendations on undergraduate medical education.* London: General Medical Council, 1993.
2. Dacre J, Nicol M, Holryod D, Ingram D. The development of a clinical skills course. *J R Coll Phys Lond* 1996; **30(4):** 318-24.
3. Goodwin J. The importance of clinical skills. *BMJ* 1995; **310:** 1281-1282.
4. Bligh J. The clinical skills unit. *Postgrad Med J* 1995; **71:** 730-732.
5. Gleeson F. Defects in postgraduate clinical skills as revealed by the objective structured long examination record (OSTLER). *Irish Med J* 1992; **85:** 11-14.
6. Johnson N, Hasler J, Toby J, Grant J. Consensus minimum standards for use in a trainer's report for summative assessment in general practice. *Br J Gen Pract* 1996; **46:** 140-144.
7. Johnson N, Hasler J, Toby J, Grant J. Content of a trainer's report for summative assessment in general practice: views of trainers. *Br J Gen Pract* 1996; **46:** 135-139.
8. Thomson A, Park KGM, Kelly DR, *et al*. Training for minor surgery in general practice - is it adequate? *J R Coll Surg Edin* 1997; **42:** 89-91.
9. Jansen JJ, Tam LH, van der Vleuten CP, *et al*. Assessment of competence in technical clinical skills in general practice. *Med Educ* 1995; **29:** 247-523.
10. Dunnington G, Wright K, Hoffman K. A pilot experience with a competency-based clinical skills assessment in a surgical clerkship. *Am J Surg* 1994; **167:** 604-607.
11. Johnson N, Hasler J, Toby J, *et al*. Pilot testing of a structured trainer's report. *Educ Gen Pract* 1997; **8:** 308-315.

Address for correspondence

Dr M H Kelly, Department of Postgraduate Medical Education, 1 Horselethill Road, Glasgow G12 9LX.