

Primary care

Model for directly assessing and improving clinical competence and performance in revalidation of clinicians

Robert K McKinley, Robin C Fraser, Richard Baker

Department of
General Practice
and Primary Health
Care, University of
Leicester, Leicester
General Hospital,
Leicester LE5 4PW

Robert K McKinley
*senior lecturer in
general practice*

Robin C Fraser
*professor of general
practice*

Richard Baker
*professor of quality in
health care*

Correspondence to:
R K McKinley
rkm@le.ac.uk

BMJ 2001;322:712-5

It is now clear that revalidation and clinical governance will drive continuing professional development in medicine in the United Kingdom.^{1 2} Thus patients, society, and the profession are to be assured that individual doctors not only are fit to practise but are providing high quality care for patients. The focus of professional revalidation is rightly moving from the requirement that practitioners merely provide evidence of participation in continuing education towards the requirement that they provide evidence that better reflects their clinical practice.^{3 4} Nevertheless, the primary screening procedures that have been proposed for revalidation are indirect (see box).⁴ If used at all, tests of clinical competence come much later in the process, but few tests include direct observation of practice. We present the case for the primacy of obtaining direct evidence of clinical competence of any doctor being revalidated; discuss the essential attributes of any process of obtaining such evidence; describe the ways in which such evidence can be gathered; explore the limitations of review tools currently available; and suggest an appropriate model for performance review.

Recent proposed components of revalidation in United Kingdom⁴

- Review of patients' case notes
- Professional values
- Patient satisfaction
- Professional relationship with patients
- Keeping up to date and monitoring performance
- Complaints procedure
- Good clinical care
- Record keeping
- Accessibility
- Team work
- Effective use of resources

Direct assessment of consultation competence

Indirect measures of competence are affected by patients and colleagues as well as by service and secular variables. High levels of patient satisfaction, for example, cannot be relied on to indicate competence, and vice versa⁵; for example, a patient may be dissatis-

Summary points

The measures currently proposed for assessing competence in clinician revalidation are mainly indirect or proxy

As the consultation is the single most important event in clinical practice, the central focus of revalidation should be the assessment of consultation competence

Such assessment should be by direct observation and satisfy five criteria—reliability, validity, acceptability, feasibility, and educational impact

Assessment of consultation competence would be followed by assessment of specific skills and regular performance review

Such an assessment procedure is recommended for use in the revalidation of all clinicians

fied with the professionally correct refusal to agree to an inappropriate request for hypnotics or antibiotics. Similarly the views of colleagues may not always truly reflect performance. For example, a doctor whose relationships with other professionals are problematic may engender negative feelings among peers but still provide good care. Furthermore, identification of poor practice through monitoring of routine data may be insensitive and inconsistent.^{6 7} Indirect review alone, therefore, is insufficient.

The cornerstone of medical practice is "the consultation . . . as all else in the practice of medicine derives from it."⁸ Accordingly, the monitoring of clinicians should focus predominantly on the direct assessment of consultation performance. Nevertheless a single demonstration of competence is not sufficient to ensure adequate performance in everyday practice—the so called competence-performance gap.⁹ Performance review can help to identify such a gap and allow its investigation and remediation. Direct assessment of competence and indirect performance review are therefore complementary, and our proposal will bolster rather than replace current UK proposals for clinical governance and revalidation. Accordingly, the

bmj.com

A table with further data on assessment tools is available on the BMJ's website

Five required attributes of an assessment process¹⁰

Reliability is a measure of the variation in scores due to differences in performance between subjects and also the correlation of assessors rating the same performance. It is generally accepted that the reliability of a regulatory assessment must be at least 0.8

Validity is the degree to which an assessment is a measure of what should be measured. Although face validity of an assessment (the extent to which an assessment measures what it purports to measure) is often discussed, this should be augmented by discussion of whether what is being assessed is what should be assessed. Validity therefore concerns both the instrument and assessment process and the challenge (cases) with which the candidate is tested. Ideally the content of the assessment should reflect the practitioner's own practice as closely as possible

Acceptability is the degree to which the assessment process is acceptable to all stakeholders. In tests of competence of a doctor the stakeholders are the doctor being assessed, the assessors, the people who provide the clinical challenge (patients or simulators), the profession, future patients of that doctor, and society

Feasibility is the degree to which the assessment can be delivered to all those who require it within real costs of staff and time constraints

Educational impact is the degree to which the assessment can assist the doctor to improve his or her performance, usually through the provision of feedback on specific strengths and weaknesses together with prioritised and specific strategies for improvement

profession can better demonstrate its commitment to establishing the competence of every practising doctor and maintaining satisfactory performance.

Required attributes of an assessment process

It is now generally accepted that any credible assessment process must have the attributes of reliability, validity, acceptability, feasibility, and educational impact (see box).¹⁰ These attributes are multiplicative—that is, if any single one is missing the overall utility of the assessment will be zero.¹⁰ Nevertheless the design of any assessment process is a compromise between these five attributes. For example, maximising reliability, validity, acceptability, and educational impact will increase costs but reduce feasibility, and vice versa. Thus, the particular emphasis given to each attribute is critically dependent on the purpose of the assessment. In formative assessment, for example, validity and educational impact are more important than high reliability, but, in any regulatory assessment to determine fitness to practise, reliability and validity are paramount. This is because of its particular importance to the doctor being assessed (who is at risk of losing his or her job), the profession (self regulation is at stake), and society (which needs the professional competence of doctors to be assured without the unnecessary loss of expensively trained professionals).

Gathering evidence of consultation competence

The assessment of consultation competence requires a judgment based on systematic observation of a practitioner's performance against validated criteria of competence. Observation can be overt or covert, live or recorded; real or simulated patients may be involved; and the assessor can be lay or professional.

Covert observation is more likely to capture the “usual” consulting behaviour of the doctor—that is, what he or she does in day to day clinical practice.¹¹ Widespread adoption of covert observation would be likely to minimise the “competence-performance gap,”⁹ but it can be ethically achieved only with prior consent from practitioners and patients—which is unlikely to be forthcoming. Consequently, any systematic programme of assessment of competence is likely to be overt.

Videotaping of consultations provides logistical advantages as the doctor and assessor(s) do not have to be in the same place at the same time. It also has potential disadvantages—for example, dependence on technical quality, unacceptability (real patients may not be expecting the examination to be videotaped), problems with validity (some patients are less likely to consent to videotaping of their consultations¹²), and difficulties in verifying physical findings. Furthermore, it cannot be emphasised strongly enough that videotaping consultations is only a means of capturing performance. It is not an assessment technique.

The clinical challenges to which the doctor is exposed can be real or simulated. Assessments based on consultations with real patients in the doctor's own place of practice have high face and content validity, but it may be difficult to regulate the difficulty or range of cases. Furthermore, particular patients may be less likely to consent to observation of their consultations. Simulated patients provide varying validity of challenges—for example, lower in disciplines in which prior knowledge of the patient is important or because of the omission of presentations (for example, of children) or of physical signs that are difficult to simulate. Simulated patients, however, allow control of the difficulty and range of challenges presented. Simulated patient encounters can also be used in different ways. They can be arranged, for example, as a series of complete consultations (“simulated surgery”)^{13 14} or as parts of consultations;¹⁵ the first option provides higher validity (but lower reliability), and the second provides higher reliability (but lower validity).¹⁰

Lastly, the assessors may be lay or professional. Any valid assessment process must, however, enable judgments to be made about the full range of required consultation competences. These range from those which lay assessors may be able to assess with little or no professional support (for example, communication and interpersonal skills) to those that demand professional input (for example, clinical problem solving and choice of clinical management options). Professional input is typically provided as checklists, but the doctor who uses idiosyncratic but still professionally appropriate methods not covered by the checklists may be unfairly penalised. Accordingly, we support the view that assessment of professionals should be performed by “professionals” but with lay input to the process and joint overview of the outcome.¹⁶

Assessment tools

Although assessment of consultation performance has been a feature of undergraduate and postgraduate clinical examinations for generations, the reliability, validity, acceptability, feasibility, and educational impact of such assessments are seldom reported. Non-standardised global assessments (traditional clinical examinations) tend to be valid but of low reliability.¹⁷ Frequently, candidates are not directly observed by the assessors, and explicit, validated criteria against which performance is to be judged are often absent. Such procedures cannot satisfy the essential five conditions.

An optimum test of consultation competence should require the observation of clinicians in complete consultations in his or her own workplace with a series of real patients (or the closest possible simulation) using an assessment tool that is structured but allows professional judgment. This implies but does not require that all assessors are professionals in the same field as the doctor being assessed.

In the United Kingdom, general practice has the longest tradition of developing assessment tools. Nevertheless, few procedures for assessing consultation competence have been specifically validated for use with established practitioners. There is also a lack of conclusive published evidence of the reliability, acceptability, feasibility, and educational impact of most assessment tools in respect of established practitioners, although “proxy” evidence is available for some (see table on the *BMJ*'s website).

Limitations

A reaccreditation process that combines assessment of consultation competence and performance review can assure the profession, its patients, and society that every practising doctor is competent in consultation skills. It will not, however, guarantee that the competent doctor puts his or her skills into practice; this requires formal review through audit and feedback from patients, which are features of the proposed annual appraisals of all medical clinicians.

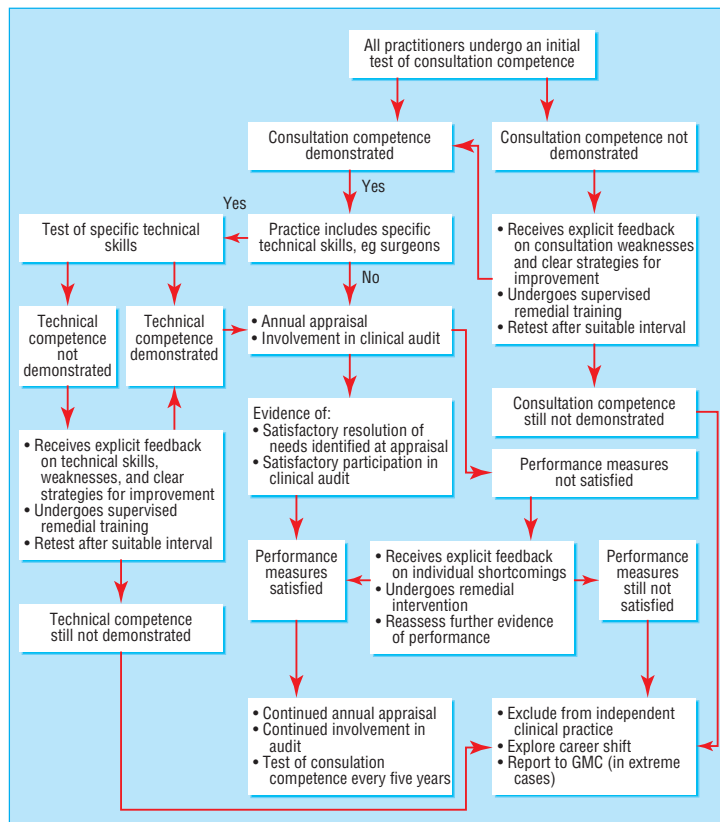
Direct assessment of consultation competence will not necessarily detect those who wilfully abuse their position of trust within the doctor-patient relationship and deliberately conceal unacceptable practice. Detection of unprofessional, negligent, or criminal behaviour will always depend on the vigilance of patients, peers, and the profession assisted by indirect assessment preferably within a formalised, regular appraisal system. Nevertheless, we believe that doctors who underperform because of lack of competence are many times more common than those who do so through malice or indifference.

Proposed model for assessing consultation competence and clinical performance

We propose an assessment model that can be applied to all clinicians (figure). At intervals, all practitioners would undergo an assessment of consultation competence that satisfies the five requirements of reliability, validity, acceptability, feasibility, and educational impact. Those who are competent would, if appropriate, undergo an additional assessment of the technical skills specific to their discipline. Doctors judged competent would enter a period of regular performance review, which would assess participation in audit, feedback from patients, peer review, complaints against them, and continuing professional development. This would result in an appraisal of their needs and an educational plan. Doctors who address their educational plan satisfactorily would continue with annual performance reviews until the process restarts.

Doctors who have not demonstrated competence in either consultation or technical skills would receive focused feedback on their weaknesses containing explicit strategies for improvement, followed by a period of supervised remedial training, after which they would be reassessed. Those who subsequently demonstrate competence would then enter annual performance review. Doctors unable to demonstrate competence would be counselled and advised to withdraw or, if necessary, removed from independent, or even all, clinical practice. Similarly, doctors unable to provide evidence of satisfactory performance and professional development at their annual performance review would also receive specific feedback on their shortcomings and undergo remedial intervention and reassessment. Those not satisfying the formal review of their performance would also be counselled to withdraw or, if necessary, removed from independent practice.

Unfortunately, no evidence base exists to help decide how long the cycle for reviewing consultation competence and performance should be, although the General Medical Council's proposed five yearly interval seems appropriate.¹⁸



Proposed model to integrate direct assessment and improvement of clinical competence with performance review and revalidation of clinicians

Thus the integration of assessment of clinical competence by direct observation of routine practice in revalidation and performance review has important advantages. By focusing on what the practitioner actually does, it enables highly context specific diagnostic evaluation, with subsequent improvement or remediation of skills. Little examination preparation is required. Paper based examinations do not test clinical competence, and, although simulations can test specific skills, ensuring validity for a particular practitioner's practice would be difficult.

We believe that such a process is feasible. Our preliminary work suggests that two trained general practitioner assessors observing a peer in a single consultation session of 10 patients achieves high levels of reliability and validity, and the assessors can provide feedback that is acceptable to practitioners (in our work the practitioners perceived that it positively influenced their practice). In addition to the costs of continuing performance review, each assessment (every five years) costs £400 per doctor (equivalent to £80 annually) plus the training costs for assessors. If assessors perform 12 assessments a year, one assessor would be required for every 30 general practitioners. We acknowledge that there are likely to be additional costs in applying the same process to hospital practitioners, especially for those with more specialised skills and consequently fewer peers. These challenges, however, are surmountable, and, even if the cost is 2.5 times that for general practitioners, it would still be only about £200 per doctor annually. There will be inevitable debate about whether these costs should be borne by the profession, employers, or purchasers. We believe, however, that for a modest investment the profession has an opportunity to show that all practitioners will both be competent in the skills required for their practice and perform subsequently to a satisfactory standard. If this opportunity is seized, medical practitioners can then rightly reclaim their position of trust having demonstrated their professional accountability and their capability of and commitment to self regulation.

Contributors: This paper condenses and synthesises extensive discussions between the authors over many years. RKMCK wrote the first draft, which was revised by RCF and RB. The figure was originally prepared by RCF and RB in response to England's chief medical officer's report *Supporting Doctors, Protecting Patients*. All authors will act as guarantors for the paper.

Funding: No special funding.

Competing interests: None declared.

- 1 General Medical Council. *Maintaining good medical practice*. London: GMC, 1998.
- 2 Donaldson L. *Supporting doctors, protecting patients*. London: Stationery Office, 1999.
- 3 Southgate L, Dauphinee D. Maintaining standards in British and Canadian medicine: the developing role of the regulatory body. *BMJ* 1998;316:697-700.
- 4 Southgate L, Pringle M. Revalidation in the United Kingdom: general principles based on experience in general practice. *BMJ* 1999;319: 1180-3.
- 5 Baker R. Pragmatic model of patient satisfaction in general practice: progress towards a theory. *Qual Health Care* 1997;6:201-4.
- 6 Frankel S, Sterne J, Smith GD. Mortality variations as a measure of general practitioner performance: implications of the Shipman case. *BMJ* 2000;320:489.
- 7 Marshall EC, Spiegelhalter DJ. Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates. *BMJ* 1998;316: 1701-4.
- 8 Spence J. The need for understanding the individual as part of the training and function of doctors and nurses. In: National Association for Mental Health, ed. *The Purpose and Practice of Medicine*. Oxford: Oxford University Press, 1960:271-80.
- 9 Rethans JJ, Sturmans F, Drop R, van der Vleuten C, Hobus P. Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice. *BMJ* 1991;303:1377-80.
- 10 Van der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. *Advances in Health Sciences Education* 1996;1:41-67.
- 11 Rethans JJ, Drop R, Sturmans F, van der Vleuten C. A method for introducing standardized (simulated) patients into general practice consultations. *Br J Gen Pract* 1991;41:94-6.
- 12 Coleman T, Manku-Scott T. Comparison of video-recorded consultations with those in which patients' consent is withheld. *Br J Gen Pract* 1998;48:971-4.
- 13 Allen J, Evans A, Foulkes J, French A. Simulated surgery in the summative assessment of general practice training: results of a trial in the Trent and Yorkshire regions. *Br J Gen Pract* 1998;48:1219-23.
- 14 Burrows PJ, Bingham L. The simulated surgery—an alternative to videotape submission for the consulting skills component of the MRCP examination: the first year's experience. *Br J Gen Pract* 1999;49:269-72.
- 15 Harden RM. What is an OSCE? *Med Teach* 1988;10:19-22.
- 16 Irvine D. The performance of doctors. I: Professionalism and self regulation in a changing world. *BMJ* 1997;314:1540-2.
- 17 Streiner DL. Global rating scales. In: Nufeld VR, Norman G, eds. *Assessing clinical competence*. New York: Springer, 1985:119-41.
- 18 General Medical Council. *Revalidating doctors: ensuring standards, securing the future*. London: GMC, 2000.

(Accepted 15 December 2000)

One hundred years ago

The future of the British Medical Association

Sir,—Five pounds a year and consent to the discipline of combination would not be much to give, if through the British Medical Association the medical profession could become organised, self-protecting, and consequently self-respecting and respectable.

In these days of organisation in all callings, the only one not doing so will naturally go to the wall, and its position become intolerable.

One thing is certain, that Conservatives, Liberals, Radicals, anti-Boers and pro-Boers, millionaires and working men, Chamberlains and Steads, are all quite agreed on one point, namely, to exploit the medical profession to the utmost of their power, and on submission, laugh in their sleeves when referring to

“the noble profession,” “the patriotism of the profession”—a profession which for the most part receives a beggarly pittance, considering the time, energy, and money spent in obtaining a degree or diploma; not to say that it is without honour in our own country, which has been practically shown twice over lately, when Ministers of the Government have held us up to derision in the House of Commons.

The position, bad as it is, will get worse so long as we remain “an unorganised mob.”

We must follow the general example; combine professionally, and take a more active share in politics—I am, etc., Centurion.

(*BMJ* 1901;i:1521)