

The accountability of clinical education: its definition and assessment

Elizabeth Murray,¹ Larry Gruppen,² Pamela Catton,³ Richard Hays⁴ & James O Woolliscroft⁵

Background Medical education is not exempt from increasing societal expectations of accountability. Competition for financial resources requires medical educators to demonstrate cost-effective educational practice; health care practitioners, the products of medical education programmes, must meet increasing standards of professionalism; the culture of evidence-based medicine demands an evaluation of the effect educational programmes have on health care and service delivery. Educators cannot demonstrate that graduates possess the required attributes, or that their programmes have the desired impact on health care without appropriate assessment tools and measures of outcome.

Objective To determine to what extent currently available assessment approaches can measure potentially relevant medical education outcomes addressing practitioner performance, health care delivery and population health, in order to highlight areas in need of research and development.

Methods Illustrative publications about desirable professional behaviour were synthesized to obtain

examples of required competencies and health outcomes. A MEDLINE search for available assessment tools and measures of health outcome was performed.

Results There are extensive tools for assessing clinical skills and knowledge. Some work has been done on the use of professional judgement for assessing professional behaviours; scholarship; and multiprofessional team working; but much more is needed. Very little literature exists on assessing group attributes of professionals, such as clinical governance, evidence-based practice and workforce allocation, and even less on examining individual patient or population health indices.

Conclusions The challenge facing medical educators is to develop new tools, many of which will rely on professional judgement, for assessing these broader competencies and outcomes.

Keywords Clinical clerkship, *standards; clinical competence, *standards; delivery of health care, *standards; public health, standards; tutor.

Medical Education 2000;34:871–879

Introduction

This paper is a product of Cambridge Conference IX, that dealt with clinical education. The paper focuses on

the need for, and implications of, introducing accountability within clinical education. However, most of what we say here applies to the whole of medical education.

¹Department of Primary Care and Population Sciences, Royal Free and University College Medical School, University College London, UK

²G1111 Towsley Center, Department of Medical Education, University of Michigan Medical School, Ann Arbor, MI, USA

³Department of Radiation Oncology, University of Toronto c/o Princess Margaret Hospital, 610 University Avenue, Toronto, ON, Canada

⁴James Cook University School of Medicine, Townsville, Queensland, Australia

⁵Medical Sciences I, University of Michigan Medical School, Ann Arbor, MI, USA

Correspondence: Elizabeth Murray, Department of Primary Care and Population Sciences, Royal Free and University College Medical School, University College London, Holborn Union Building, Highgate Hill, London N19 3UA, UK

Why should medical educators be accountable?

As medical educators, we have a responsibility to the community we serve.^{1,2} Our educational programmes should produce graduates with the characteristics our patients require. New curricula are emerging internationally which aim to produce doctors who are not only clinically competent, but also excel in interpersonal skills, team working, judicious use of resources, professional behaviours such as ethical practice and altruism, and are capable of adapting in response to changing societal expectations and developing medical science.

There are three factors driving the move toward greater accountability in medical education. The first is that medical education is expensive. The Service Increment for Teaching (SIFT) budget, designed to cover the additional service costs of teaching undergraduate medical students in clinical settings in England and Wales was £479 million in 1998–99.³ In an environment where there is keen competition for resources available for health care, medical educators must demonstrate that teaching is provided in the most cost-effective way.⁴ This requires methods of assessing effectiveness, i.e. whether learners have learnt what is intended. Moreover, she who pays the piper calls the tune; and funders have a reasonable expectation that medical education will address their priorities, be they governmental or private.

Secondly, society requires accountability from the medical profession. Throughout the Western world, patients, governments and insurers, are calling for doctors to practice up-to-date, cost-effective medicine in a humanistic way. The profession is beginning to respond to these societal expectations by developing codes of behaviour governing doctors' professional and personal behaviours.^{5–7} Statutory bodies such as the General Medical Council in the UK have led the way, with detailed descriptions of the standards expected. Part of professional accountability includes self-regulation of educational standards; unless medical educators are able to document responsiveness to societal expectations through delivery of educational programmes, standards may be imposed by external regulatory bodies.

Finally, part of the rationale for investment in medical education is the belief that it has an effect on service delivery and health care. Funding may be provided for educational programmes with specific aims, such as improving health care in under-served areas, or increasing the proportion of graduates who wish to practice in primary care. Funders have a right to evidence on how well the programme achieves these

aims. Internationally, there is increasing emphasis on evidence-based medicine. In the UK, the Health Technology Assessment Research and Development Programme has been established to ensure that all new technology is rigorously evaluated before it is introduced into the National Health Service.⁸ Institutions such as the National Institute for Clinical Excellence have been established to review all aspects of clinical practice to ensure that patients receive the most cost-effective treatments available.⁸ In this climate medical educators have a duty to provide evidence that their intervention does indeed have the desired effect on service delivery and health care.

To whom are medical educators accountable?

Medical educators are accountable to all those with a stake in the outcomes of medical education at all levels: undergraduate, graduate and continuing professional development (the UGC continuum). They include: funders of education, e.g. governments, purchasers of health care such as insurers and health authorities, self-funding students (or their parents), and tax-payers; statutory bodies with responsibility for professional self-regulation, e.g. General Medical Council, Association of American Medical Colleges (AAMC); patients; and learners (undergraduates, residents and doctors participating in continuing professional development).

What should medical educators be accountable for?

Many of the stakeholders described above have made clear demands on the profession in terms of the attributes they expect of medical practitioners. The most detailed description of the professional and personal behaviours expected of all registered doctors has been developed by the General Medical Council of the UK.⁵ It details the standard of financial probity, ethical

Key learning points

- Medical educators must become more accountable as:
 - 1 competition for resources requires us to demonstrate sound stewardship of educational expenditure
 - 2 society requires greater professional accountability across the whole of medicine
 - 3 the culture of evidence-based medicine requires us to demonstrate that education has the desired effects on health care and service delivery.
- Accountability requires assessment tools across a wide range of attributes and health outcomes, including individual and group attributes of physicians and health indices of individual patients and populations.
- Currently, most assessment tools concentrate on knowledge and clinical skills: there is a major research and development agenda required to develop new assessment tools across the complete range of attributes society requires.

behaviour, trustworthiness and respect for patients and colleagues expected of all doctors, irrespective of their discipline or seniority. This is the only code of behaviour which spells out the collective responsibility of the profession, and lays a duty on all physicians to take appropriate action to protect patients from poorly performing colleagues.

Very importantly, the GMC has also invested in developing procedures for identifying and dealing with poorly performing doctors, and has expended considerable time and energy in developing reliable and valid procedures for examining an individual doctor's practice.^{9,10} This is an example of the level of accountability that funders and patients are expecting. No longer are statements of desired outcomes sufficient. Demonstrating that the desired outcomes are being achieved is the expectation for medical educators. Moreover, the emphasis on broader outcomes than the traditional clinical skills and knowledge has substantial implications for medical education across the UGC continuum.

Educators have responded to these initiatives by specifying competencies that students must acquire before progression. At the Medical School level, the Association of American Medical College's Medical School Objectives Project (MSOP)¹¹ developed four competencies (altruism; knowledgeable; skilful; dutiful) that US medical schools should assure all graduating medical students possess. For postgraduate education, the Royal Colleges of Physicians and Surgeons in Canada delineated the roles that a specialist needed to fulfil, and the implications of these for the education of residents.⁶ Similarly, at the residency level, the Accreditation Council for Graduate Medical Education (ACGME), the overarching accrediting body for Graduate Medical Education in the United States, has recently developed six competencies (patient care; medical knowledge; practice-based learning and improvement; interpersonal and communication skills; professionalism; and systems-based practice) for which residency training programmes will be responsible. Residency Review Committees, the groups responsible for developing the accrediting standards for each discipline, are now in the process of developing standards that are based on these competencies for all of the specialty and sub-specialty programmes.⁷

Societal expectations extend beyond the individual standards of the doctors in the system. Health care purchasers, whether state or private, are interested in documented improvements in health outcomes, both at the individual and population level. Once again, the United States has led the way, with the National Committee on Quality Assessment's (NCQA) Health

Employers Data Information Service (HEDIS).¹² This is a databank of health outcome measures, ranging from the percentage of diabetic patients with normal glycosylated haemoglobins, to the stage at which breast cancer is identified, to paediatric and adult immunization rates and prescribing rates for key conditions such as beta-blockers after myocardial infarctions. Health Maintenance Organizations and insurers require the routine provision of this information from providers and use it as a quality marker. This emphasis on health outcomes, particularly at population level, poses a particular challenge for medical educators, who have traditionally focused on what the individual learner can do, rather than their performance in practice. But the purpose of medical education at all levels is to improve health care and health outcomes – and it is against outcomes like this that the products of an educational system should and will be measured. Thus, if a medical school has a mission to produce doctors with an interest in primary care who practise in under-served areas, it should be evaluated in terms of its success in achieving this. Similarly, if an educational programme is designed to create life-long learners who practice evidence-based health care, then its graduates should be evaluated against this outcome, notwithstanding the other determinants of health outcomes, such as career structure, remuneration systems and infrastructure.

How can medical educators become more accountable?

Given that assessment drives learning, one challenge facing educators today is to find robust and feasible tools of assessing these broader competencies for use across the UGC continuum. The aim of this paper is to provide examples of the competencies educators need to assess, and examine the assessment tools currently available with the intention of highlighting areas in particular need of research and development. We concentrated on assessment (rather than curriculum or instructional methods) as the educational implications of these outcomes will have to be addressed locally, and it will not be possible to evaluate the outcome of educational initiatives without valid, reliable and feasible assessment tools.

Methodology

As a first step in this process we examined the publications about desirable professional behaviour from the USA,⁷ Canada,⁶ UK⁵ and Australia.^{13,14} These countries were chosen as a convenient sample of those

showing leadership in this area to illustrate the major emerging themes of professional accountability.

These documents identified numerous competencies and outcomes that highlighted the breadth of perspectives of various stakeholders and emphasized that the attributes of individual physicians is only one part of the equation. We have therefore grouped potential outcomes of medical education into individual practitioner attributes, attributes of the profession as a group, and impact on individual patients and population health. (Table 1).

We conducted a series of searches of MEDLINE for references that described methods and instruments for assessing these various outcomes. The searches utilized the terms in column one of Tables 2–5 as general search terms, that is, not as medical subject heading (MESH) terms. This search strategy returned references that included the search term anywhere in the title or abstract of the record. The results of the searches were then examined to identify examples of research that either developed or examined assessment methods for those outcomes. The searches were by no means exhaustive, but were intended as an initial summary of helpful points of departure for further efforts to develop and apply reliable and valid assessment methods. These searches identified illustrative assessment tools for the components of many of the cells, but it rapidly became clear that the assessment of some outcomes is still in a very rudimentary stage of

development. Tables 2–5 also provide some commentary on the validity and reliability of these assessments, with the aim of highlighting areas for development and research in the field of assessment.

Results

Table 2 presents some assessment tools available for attributes of individual practitioners. These are relatively well addressed, with extensive tools available for assessment of clinical skills and knowledge. However more work is clearly needed in assessing attributes such as professional behaviours (altruism, honesty, integrity, responsibility, respect for patients and colleagues, confidentiality and ethical behaviour); efficiency; cross-cultural competency; multiprofessional team working; and scholarship. Table 3 examines the attributes of practitioners in groups, such as clinical governance, evidence-based practice and workforce allocation. This is a complex area, which requires considerably more work on developing the underlying concepts, before assessment tools can be developed. Tables 4 and 5 examine outcomes of medical education from the point of view of individual patients and a population perspective. The literature here largely comes from clinical and health-services research; adapting some of these measures for widespread use as an outcome measure for medical education is likely to be problematic, and will require considerable cross-disciplinary collaboration.

| | Doctor | Patient |
|-----------------------|--|--|
| Individual attributes | clinical skills professional behaviours: altruism honesty and integrity responsibility respect for patients and colleagues confidentiality ethical behaviour efficiency cross-cultural competency multiprofessional team working scholar | improved health patient enablement patient satisfaction |
| Group attributes | clinical governance evidence-based practice workforce allocation | reduced morbidity/mortality reduced iatrogenic illness accessibility of services equity of resource allocation culturally appropriate services |

Table 1 An organizational framework for potential outcomes of medical education

This table is a synthesis of currently available documents on desirable outcomes of medical education such as the GMC 'Good Medical Practice',⁵ CanMEDS 2000,⁶ ACGME⁷ and the Australian Medical Association publications.^{13,14} It is intended to provide illustrative examples of desirable outcomes.

Table 2 Individual practitioner

| Attribute | Assessment tools available | validity/reliability/feasibility/comments |
|---|---|---|
| <p>Clinical skills: diagnostic competence including history taking and physical examination technical competence communication skills</p> | <p>OSCE simulations long and short cases simulated patients written (MEQ, MCQ, PMP, etc.) oral observation (immediate or videotaped) log books</p> | <p>an extensive literature is available on the validity and reliability of these methods^{15,16}</p> |
| <p>Professional behaviours: altruism honesty and integrity responsibility respect for patients and colleagues confidentiality ethical behaviour</p> | <p>overall: rating scales with various informants (peers, patients, supervisors), questionnaires value conflict methodology, concern/compliment cards (critical incidents),¹⁷ clinical ratings</p> | <p>definitions of desirable attitudes in behavioural terms remains one of the major challenges in this domain. Reliability of 'trait' expression in situations is low. Stability of values over time is also unknown. Ideal assessment would be: behaviourally based; sampled over time, context, observers; mixture of planned and opportunistic samples (to catch the chameleons); have multiple methods (ratings, incident reports, OSCEs); hence would be expensive and time consuming. Self-assessment necessary for motivating change</p> |
| <p>Efficiency</p> | <p>Relative value unit system in radiology¹⁸</p> | <p>efficiency includes cost-effective use of resources (time, investigations, treatments) need quality markers to avoid simplistic measures of increased throughput¹⁹ likely to be highly context dependent²⁰</p> |
| <p>Cross-cultural competency</p> | <p>simulations, OSCEs</p> | <p>to date, work has concentrated on teaching this competency rather than assessing it^{21,22}</p> |
| <p>Multiprofessional team working</p> | <p>observation of practice rating scales by team members^{23,24} simulations of team problem situations (e.g. conflict)</p> | <p>rating scales and criteria for assessment needed need to identify and define component parts including process (e.g. conflict management, negotiation communication, planning, co-operation) and outcome (e.g. patient care, practice management)</p> |
| <p>Scholar develop, implement and monitor a personal continuing education strategy</p> | <p>OSCE/other simulations; satisfactory performance requires admitting ignorance self assessment</p> | <p>self-assessment found to be inaccurate for both undergraduates and practising clinicians²⁶⁻²⁸</p> |
| <p>critically appraise sources of medical information</p> | <p>triple jump examination (TJE), clinical simulations, performance of EBM component skills (question formulation, evidence assemblage, critical appraisal, application), physician congruence with guidelines vignette-based measures²⁵</p> | <p>TJE has low reliability and is time consuming,¹⁵ insufficient data on other methods currently available</p> |
| <p>facilitate learning of patients, house staff/students and other health professionals contribute to development of new knowledge</p> | <p>fiction of indicators such as contact hours, student satisfaction, peer observation, etc. traditional output measures for academics (number of publications, grant funding, etc.)</p> | <p>little data on validity/reliability measures needed for those not in academic settings</p> |

| Attribute | Assessment tools | Validity/reliability/feasibility/comments |
|---|--|---|
| clinical governance evidence-based practice | guideline adherence/ utilization | emerging concept ²⁹ validity depends on the quality of the guidelines ³⁰ and the data used to define adherence reliability is unknown quality of clinical data and IT availability are major determinants of feasibility Primary care groups in the UK having to address this with the Health Improvement Programme ³¹ |
| workforce allocation | workforce statistics specialty location under-served populations | depends on the quality of the infrastructure and data |

Table 3 Practitioners in groups**Table 4** Individual patients

| Attribute | Assessment tool | Validity, reliability, feasibility |
|-----------------------------|--|--|
| health status/ morbidity | quality of life measures, e.g. SF-36 ³² Euroqol ³³ pain scales individual rating scales for specific conditions | psychometric properties of most of these scales have been well described feasibility of widespread use is questionable |
| patient enablement | Patient Enablement Instrument ³⁴ Decisional Conflict Scale ³⁵ disease specific instruments, e.g. Diabetes Empowerment Scale ³⁶ | developed for use in primary care, validated in Scotland, no data on use in secondary care a research tool developed for use with shared decision making programmes; no data available on routine use variety of disease specific instruments available with extensive validation data |
| patient satisfaction | patient surveys ³⁷ | feasibility and affordability of widespread/repeated use is questionable |

Discussion

The tables demonstrate the extent to which research on assessment tools in medical education has concentrated on attributes of the individual practitioner, particularly in the domains of knowledge and clinical skills. As assessment drives learning, learners are unlikely to concentrate on acquiring the broader competencies and health outcomes required by society unless the assessment tools used examine these attributes.

In order to achieve this desired educational impact, assessment tools may have to move towards more qualitative measures, which rely on professional judge-

ments.⁴³ Validity and reliability are likely to come from multiple sampling and triangulation of data. Examples of the successful implementation of this approach include the introduction of summative assessment for all general practice registrars in the UK. Competencies tested go beyond clinical skills and knowledge, and include effective communication, adequate consultation skills and the ability to appraise the practitioners own working practices and institute change as appropriate. Assessment is by multiple choice questions, a trainers report and submission by the candidate of a video of consultations and a written practice-based audit which are judged by trained GP assessors.⁴⁴

Table 5 Patients in groups:

| Attribute | Assessment tool | Validity/reliability/feasibility/comments |
|---|---|--|
| mortality rates | | |
| morbidity | | |
| improved health status | HEDIS data | difficult to tease apart 'system' problems and identify specific causes and remedies |
| quality of life | | |
| decreased iatrogenic illness | quality of life measures, audits | partnerships between educators and managers will be essential; |
| reduced incidence and prevalence of preventable disease | population samples, case note reviews, public health measures ³⁸ | educators need to acquaint themselves with the Health Service management literature |
| appropriate use of technology and resources | utilization review ³⁹⁻⁴¹ | validity is very uncertain |
| accessibility of services | HEDIS data | |
| | aspect of patient satisfaction measures ^{37,42} | |
| equity of resource allocation | utilization review | validity is very uncertain |
| culturally appropriate services | aspect of patient satisfaction measures | |

Southgate *et al.*, working on the performance procedures for poorly performing doctors, have devoted considerable thought to developing a valid and reliable method of assessing a doctor's performance in practice. The definition of 'acceptable', 'cause for concern', or 'unacceptable' is overtly based on the judgement of the assessors (both lay and medical) on the degree to which the doctor's performance measures up to the standards reasonably expected of any other competent doctor, in the areas detailed in good medical practice. These measures include a portfolio of approaches, including examination of medical records; case-based discussions; observation of consultations; tour of the doctors working environment; up to 20 structured interviews with third parties including patients, colleagues, health service managers, and nurses; and an interview with the doctor under review. As the stakes are high, namely the potential removal of a doctor's licence to practice, the validity and reliability of these performance procedures must be robust enough to withstand legal challenge. The time and cost of gathering sufficient data in this manner limits the use of these procedures to doctors identified as at high risk of poor performance through other, less rigorous but more feasible methods, such as local peer review.⁹

These examples show that assessment of individual practitioners across a broad range of competencies is feasible, although much more research and development is needed. More problematic is assessment of attributes of groups of practitioners. One example of this approach is the introduction of clinical governance to primary care groups in the UK; as yet there is little

data on the acknowledged difficulties or potential outcomes of this approach.

Health service managers have had more experience than medical educators of examining health outcomes, both at individual patient and at population levels. Use of these types of data will not be straightforward: on the one hand it will be difficult to tease apart 'system' problems to identify specific causes and remedies and on the other hand there are difficulties in determining the relative impact of educational input compared to other external changes.

Clearly, the availability of assessment tools and measures of health outcome, is only one small part of demonstrating the accountability of medical education in terms of professional competencies and effect on service delivery and health care. Not only will routine assessment procedures have to change dramatically to include such measures, but also there will have to be conscious linkage of educational programmes to measurable patient outcomes. Only when the effectiveness of different educational programmes can be determined, will we be in a position to consider cost-effectiveness, and hence meet one of the major agendas of funders of medical education. These are huge challenges facing medical educators; but unless we rise to the challenge, we may lose our ability to influence future educational developments.

Acknowledgements

We are grateful to Sue Elliott for her contribution to the discussion and development of the original concept, and to the referees for their helpful suggestions.

Contributors

All authors contributed equally to the original concept and the development of the paper during the Ninth Cambridge Conference. E.M. and L.G. wrote the first draft of the paper, all authors contributed to subsequent drafts; E.M. co-ordinated further authorship and is guarantor for the paper.

References

- Boelen C, Heck JE. *Defining and Measuring the Social Accountability of Medical Schools*. Geneva: World Health Organization; 1995.
- Ogrinc GS, Headrick LA, Boex JR and the Value of Education Research Group. Understanding the value added to clinical care by educational activities. *Acad Med* 1999;74:1080–6.
- The Government's Expenditure Plans. 1998–99. <http://www.official-documents.co.uk/document/cm39/3912/3912f4-7.htm>.
- Mennin SP. Standards for teaching in medical schools: double or nothing. *Med Teach* 1999;21:543–5.
- General Medical Council. *Good Medical Practice*. London: GMC. 1999. www.gmc-uk.org.
- The Royal College of Physicians and Surgeons of Canada's Canadian Medical Education Directions for Specialists Project. *Skills for the New Millennium: Report of the Societal Needs Working Group (CanMEDS 2000 Project)*, Ottawa 1996. 2000.
- Accreditation Council for Graduate Medical Education. <http://www.acgme.org/outcomes/comptv13.htm>.
- <http://www.hta.nhsweb.nhs.uk/>
- Southgate L. Assessing professional performance. In: Kent A, ed. *Proceedings of the 9th International Ottawa Conference on Medical Education*, Cape Town, 1–3 March. Cape Town: Conference Management Centre, UCT Medical School; 2000; pp. 152.
- Southgate L, Pringle M. Revalidation in the United Kingdom: general principles based on experience in general practice. *BMJ* 1999;319:1180–3.
- The Medical School Objectives Writing Group. Learning objectives for medical student education – guidelines for medical schools. Report I of the Medical Schools Objectives Project. *Acad Med* 1999;74:13–8.
- www.ncqa.org
- Australian Medical Association. www.ama.org.au
- Hays RB, Miller G, Booth B, Harris M, Harris B, Stirton J. The development of general practice standards for Australia. *Med Educ* 1998;32 (2):199–204.
- Jolly B, Grant J, eds. *The Good Assessment Guide*. London: Joint Centre for Education in Medicine; 1997.
- Newble D, Jolly B, Wakeford R, eds. *The Certification and Re-certification of Doctors. Issues in the Assessment of Clinical Competence*. Cambridge: Cambridge University Press; 1994.
- Papadakis M, Osborn EHS, Cooke M, et al. A strategy for the detection and evaluation of unprofessional behaviour in medical students. *Acad Med* 1999;74:980–91.
- Conoley PM, Vernon SW. Productivity of radiologists: estimates based on analysis of relative value units. *Am J Roentgenol* 1991;157:1337–40.
- Camasso MJ, Camasso AE. Practitioner productivity and the product content of medical care in publicly supported health centres. *Soc Sci Med* 1994;38:733–48.
- Seller RH, Lobley M. Efficient diagnosis of common complaints: a comparative study in the United States and England. *J Fam Pract* 1991;33:41–6.
- Borkan JM, Neher JO. A developmental model of ethnosensitivity in family practice training. *Fam Med* 1991;23:212–7.
- Kai J, ed. *Valuing Diversity: a Resource for Effective Health Care of Ethnically Diverse Communities*. London: Royal College of General Practitioners; 1999.
- Heinemann GD, Schmitt MH, Farrell MP, Brallier SA. Development of an attitude toward health care teams scale. *Eval Health Prof* 1999;22:123–42.
- Baggs J. Collaboration and satisfaction about care decisions scale. Development of an instrument to measure collaboration and satisfaction about care decisions. *J Adv Nursing* 1994;20:176–82.
- Holloway RG, Gifford DR, Frankel MR, Vickrey BG. A randomised trial to implement practice recommendations: designs and methods of the Dementia Care Study. *Control Clin Trials* 1999;20:369–85.
- Jankowki J, Crombie I, Block R, Mayet J, McLay J, Struthers AD. Self-assessment of medical knowledge: do physicians overestimate or underestimate? *J Roy Coll Physicians London* 1991;25:306–8.
- Woolliscroft JO, TenHaken J, Smith J, Calhoun JG. Medical students clinical self-assessments: comparisons with external measures of performance and the students' self-assessments of overall performance and effort. *Acad Med* 1993;68:285.
- Jansen JJ, Grol RP, Crebolder HF, Rethans J, van der Vleuten CP. Failure of feedback to enhance self-assessment skills of GPs. *Teaching Learning Medicine* 1998;10:145–51.
- Department of Health. Clinical governance in the new NHS. HSC99/065 (www.doh.gov.uk/coinh.htm).
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based Medicine. How to Practice and Teach EBM*. New York: Churchill Livingstone; 1997.
- Roland M, Baker R. *Clinical governance: a practical guide for primary care teams*. Manchester: University of Manchester; 1999.
- Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160–4.
- Euroqol Group. Euroqol: a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16: 199–208.
- Howie JGR, Heaney DJ, Maxwell M, Walker JJ, Freeman GK, Rai H. Quality at general practice consultations: cross sectional survey. *BMJ* 1999;319:738–43.
- O'Connor AM. Validations of a decisional conflict scale. *Med Decis Making* 1995;15:25–30.

- 36 Anderson RM, Fitzgerald JT, Funnell MM. The diabetes empowerment scale (DES): a measure of psychosocial self-efficacy. *Diabetes Care* 2000;**23**: in press.
- 37 Consumers Health Forum. *Integrating Consumer Views About Quality in General Practice*. Canberra: Commonwealth of Australia; 1995.
- 38 Wilson RM, Harrison BT, Gibberd RW, Hamilton JD. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. *Med J Australia* 1999;**170**(9):411–5.
- 39 Werneke U, MacFaul R. Evaluation of appropriateness of paediatric admission. *Arch Dis Childhood* 1996;**74**(3): 268–73.
- 40 Restuccia JD. The evolution of hospital utilization review methods in the United States. *Int J Quality Health Care* 1995;**7** (3):253–60.
- 41 Ash A. The design and analysis of hospital utilization studies. *Int J Quality Health Care* 1995;**7** (3):245–52.
- 42 Baker R. The reliability and criterion validity of a measure of patients' satisfaction with their general practice. *Fam Pract J* 1991;**8** (2):171–7.
- 43 Van der Vleuten C. www.educ.unimaas.nl/ottawa.
- 44 Conference of Postgraduate Advisers in General Practice Universities of United Kingdom (UKCRA). *Summative Assessment for General Practice Training*. Hants: National Office for Summative Assessment; 1998.

Received 13 April 2000; editorial comments to authors 5 May 2000; accepted for publication 7 June 2000