ABC of learning and teaching in medicine
Evaluation

Jill Morrison

Evaluation is an essential part of the educational process. The focus of evaluation is on local quality improvement and is analogous to clinical audit. Medical schools require evaluation as part of their quality assurance procedures, but the value of evaluation is much greater than the provision of simple audit information. It provides evidence of how well students' learning objectives are being achieved and whether teaching standards are being maintained. Importantly, it also enables the curriculum to evolve. A medical curriculum should constantly develop in response to the needs of students, institutions, and society. Evaluation can check that the curriculum is evolving in the desired way. It should be viewed positively as contributing to the academic development of an institution and its members.

Evaluation versus research

Evaluation and educational research are similar activities but with important differences. Research is usually aimed at producing generalisable results that can be published in peer reviewed literature, and it requires ethical and other safeguards. Evaluation is generally carried out for local use and does not usually require ethics committee approval. Evaluation has to be carefully considered by curriculum committees, however, to ensure that it is being carried out ethically. Finally, evaluation is a continuous process, whereas research may not become continuous if the answer to the question is found.

What should be evaluated

Evaluation may cover the process and/or outcome of any aspect of education, including the delivery and content of teaching. Questions about delivery may relate to organisation—for example, administrative arrangements, physical environment, and teaching methods. Information may also be sought about the aptitude of the teacher(s) involved. The content may be evaluated for its level (it should not be too easy or too difficult), its relevance to curriculum objectives, and integration with previous learning.

Outcome measures may show the impact of the curriculum on the knowledge, skills, attitudes, and behaviour of students. Kirkpatrick described four levels on which to focus evaluation; these have recently been adapted for use in health education evaluation by Barr and colleagues. Some indication of these attributes may be obtained by specific methods of inquiry—for example, by analysing data from student assessments.

Evaluation in curriculum planning

Evaluation should be designed at the start of developing a curriculum, not added as an afterthought. When an educational need has been identified, the first stage is to define the learning outcomes for the curriculum. The goals of the evaluation should be clearly articulated and linked to the outcomes.

Clarifying the goals of evaluation will help to specify the evidence needed to determine success or failure of the training. A protocol should then be prepared so that individual responsibilities are clearly outlined.

Purpose of evaluation
- To ensure teaching is meeting students' learning needs
- To identify areas where teaching can be improved
- To inform the allocation of faculty resources
- To provide feedback and encouragement for teachers
- To support applications for promotion by teachers
- To identify and articulate what is valued by medical schools
- To facilitate development of the curriculum
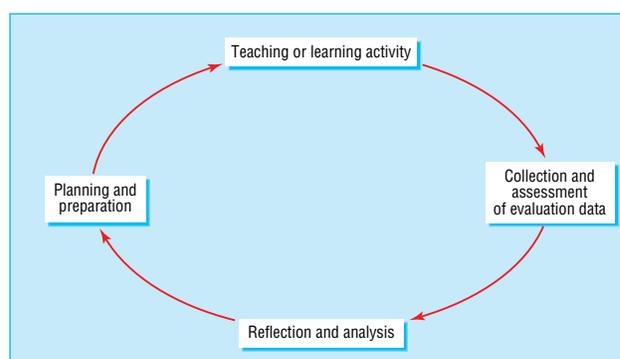
Evaluation cycle. From Wilkes et al (see "Further reading" box)

Kirkpatrick's four levels on which to focus evaluation*
- Level 1—Learner's reactions
- Level 2a—Modification of attitudes and perceptions
- Level 2b—Acquisition of knowledge and skills
- Level 3—Change in behaviour
- Level 4a—Change in organisational practice
- Level 4b—Benefits to patients or clients

*Adapted by Barr et al (see "Further reading" box)

The full impact of the curriculum may not be known until some time after the student has graduated

Questions to ask when planning an evaluation
- What are the goals of the evaluation?
- From whom and in what form will data be collected?
- Who will collect and analyse data?
- What type of analysis, interpretation, and decision rules will be used and by whom?
- Who will see the results of the evaluation?

Designing evaluation

An ideal evaluation method would be reliable, valid, acceptable, and inexpensive. Unfortunately, ideal methods for evaluating teaching in medical schools are scarce.

Establishing the reliability and validity of instruments and methods of evaluation can take many years and be costly. Testing and retesting of instruments to establish their psychometric properties without any additional benefit for students or teachers is unlikely to be popular with them. There is a need for robust "off the shelf" instruments that can be used to evaluate curriculums reliably. The process of evaluation itself may produce a positive educational impact if it emphasises those elements that are considered valuable and important by medical schools.

Participation by students

Several issues should be considered before designing an evaluation that collects information from students.

Competence—Students can be a reliable and valid source of information. They are uniquely aware of what they can consume, and they observe teaching daily. They are also an inexpensive resource. Daily contact, however, does not mean that students are skilled in evaluation. Evaluation by students should be limited to areas in which they are competent to judge.

Ownership—Students who are not committed to an evaluation may provide poor information. They need to feel ownership for an evaluation by participating in its development. The importance of obtaining the information and the type of information needed must be explicit. Usually the results of an evaluation will affect only subsequent cohorts of students, so current students must be convinced of the value of providing data.

Sampling—Students need to feel that their time is respected. If they are asked to fill out endless forms they will resent the waste of their time. If they become bored by tedious repetition, the reliability of the data will deteriorate. One solution is to use different sampling strategies for evaluating different elements of a curriculum. If reliable information can be obtained from 100 students, why collect data from 300?

Anonymity is commonly advocated as a guard against bias when information is collected from students. However, those who support asking students to sign evaluation forms say that this helps to create a climate of responsible peer review. If students are identifiable from the information they provide, this must not affect their progress. Data should be collected centrally and students' names removed so that they cannot be identified by teachers whom they have criticised.

Feedback—Students need to know that their opinions are valued, so they should be told of the results of the evaluation and given details of the resulting action.

Methods of evaluation

Evaluation may involve subjective and objective measures and qualitative and quantitative approaches. The resources devoted to evaluation should reflect its importance, but excessive data collection should be avoided. A good system should be easy to administer and use information that is readily available.

Interviews—Individual interviews with students are useful if the information is sensitive—for example, when a teacher has received poor ratings from students, and the reasons are not clear. A group interview can provide detailed views from students or teachers. A teaching session can end with reflection by the group.
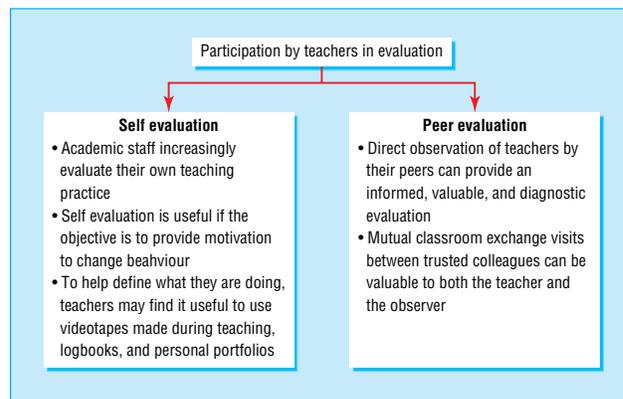

Characteristics of an ideal evaluation

- Reliability
- Validity
- Acceptability—to evaluator and to person being evaluated
- Inexpensiveness

---

To reduce possible bias in evaluation, collect views from more than one group of people—for example, students, teachers, other clinicians, and patients

---

Areas of competence of students to evaluate teaching and curriculum

- Design: whether the curriculum enables students to reach their learning objectives; whether it fits well with other parts of the curriculum
- Delivery: attributes of teacher and methods used
- Administrative arrangements

Participation by teachers in evaluation

Self evaluation
- Academic staff increasingly evaluate their own teaching practice
- Self evaluation is useful if the objective is to provide motivation to change beahviour
- To help define what they are doing, teachers may find it useful to use videotapes made during teaching, logbooks, and personal portfolios

Peer evaluation
- Direct observation of teachers by their peers can provide an informed, valuable, and diagnostic evaluation
- Mutual classroom exchange visits between trusted colleagues can be valuable to both the teacher and the observer

Issues relating to students' participation in evaluation may also apply to teachers, but self evaluation and peer evaluation are also relevant

I THINK THAT'S YOU'LL FIND ... BUT IN VALUATION! THIS CASE, NOT FAR OFF!
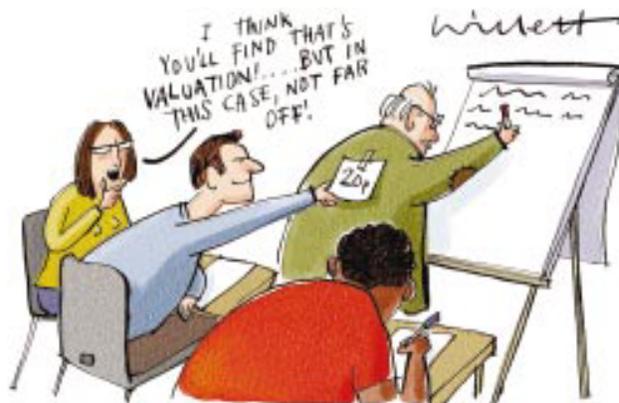
Surveys—Questionnaires are useful for obtaining information from large numbers of students or teachers about the teaching process. Electronic methods for administering questionnaires may improve response rates. The quality of the data, however, is only as good as the questions asked, and the data may not provide the reasons for a poorly rated session.

Information from student assessment—Data from assessment are useful for finding out if students have achieved the learning outcomes of a curriculum. A downward trend in examination results over several cohorts of students may indicate a deficiency in the curriculum. Caution is needed when interpreting this source of information, as students' examination performance depends as much on their application, ability, and motivation as on the teaching.

Completing the evaluation cycle

The main purpose of evaluation is to inform curriculum development. No curriculum is perfect in design and delivery. If the results of an evaluation show that no further development is needed, doubt is cast on the methods of evaluation or the interpretation of the results.

This does not mean that curriculums should be in a constant state of change, but that the results of evaluation to correct deficiencies are acted on, that methods continue to improve, and that content is updated. Then the process starts all over again.

Worked example

Background

Clinical teaching staff think that students are becoming weaker at examining cranial nerves. The examination scores for that part of the objective structured clinical examination (OSCE) carried out at the end of year three show a decline over several years. Three focus groups are held with students in year four, and several clinical teachers are interviewed. The results suggest that the decline is due to fewer appropriate patients presenting at outpatient sessions where cranial nerve examination is taught and to a lack of opportunities for practising examination skills.

Intervention

A teaching session is designed for delivery in the clinical skills centre. After that, students should be able to do a systematic examination of cranial nerves. They should also recognise normal signs and know which common abnormalities to look for. Sessions are timetabled for practising skills learnt during the teaching session.

Evaluation

A questionnaire is developed for completion by a third of students. It seeks their views on the teaching process, including the teaching skills of the tutor, physical aspects of the teaching environment, appropriateness of the teaching material, and opportunities for practising examination skills. Outcome measures include comparison of examination scores for students in the previous cohort with those participating in the teaching session, plus a questionnaire for all clinical supervisors for neurology in the following year to get their views about students' examination skills. A tenth of students with a range of scores in the relevant part of the OSCE are interviewed to find out the reasons for their varied scores. The evaluation results are disseminated widely to staff and students.

Jill Morrison is professor of general practice and deputy associate dean for education at Glasgow University.

> Questionnaire surveys are the most common evaluation tool

OSCE = objective structured clinical examination

Examples of methods of evaluation

Key points

Evaluation should:
- Enable strategic development of a curriculum
- Be a positive process that contributes to the academic development of a medical school

The goals of an evaluation should:
- Be clearly articulated
- Be linked to the outcomes of the teaching

When carrying out an evaluation:
- More than one source and type of information should be sought
- The results should be fed back to participants and details of the resulting action given

Learners need:
- To be involved in developing an evaluation
- To feel their time is respected
- To know their opinions are valued and acted on

Evaluators must:
- Act on the results of the evaluation to correct deficiencies, improve methods, and update content
- Repeat the process

Further reading
- Robson C. *Small scale evaluation.* London: Sage, 2000.
- Cohen L, Manion L. *Research methods in education.* 4th ed. London: Routledge, 1994.
- Snell L, Tallett S, Haist S, Hays R, Norcini J, Prince K, et al. A review of the evaluation of clinical teaching: new perspectives and challenges. *Med Educ* 2000;34:862-70.
- Barr H, Freeth D, Hammick M, Koppel, Reeves S. *Evaluations of interprofessional education: a United Kingdom review of health and social care.* London: CAIPE/BERA, 2000.
- Wilkes M, Bligh J. Evaluating educational interventions. *BMJ* 1999;318:1269-72.

The ABC of learning and teaching in medicine is edited by Peter Cantillon, senior lecturer in medical informatics and medical education, National University of Ireland, Galway, Republic of Ireland; Linda Hutchinson, director of education and workforce development and consultant paediatrician, University Hospital Lewisham; and Diana F Wood, deputy dean for education and consultant endocrinologist, Barts and the London, Queen Mary's School of Medicine and Dentistry, Queen Mary, University of London. The series will be published as a book in late spring.