

Improving clinical education through evaluation

H. A. P. WOLFHAGEN, W. H. GIJSELAERS, D. DOLMANS, G. ESSED & H. G. SCHMIDT

Department of Educational Development and Research, University of Maastricht, The Netherlands

SUMMARY *Although clinical clerkships are an important and essential part of any medical training program, medical schools' policy makers seem to exert little control on the educational effectiveness of clerkships. However, educational quality of clerkships may be assured and improved by using information from evaluation. The success of the usefulness of evaluation results is clearly associated with the fulfilment of a number of conditions: willingness to adopt a critical attitude, willingness to analyse the existing situation, opportunities to discuss and carry on a dialogue, availability of a plan of action and continuous collection of evaluative data. This paper describes these conditions and the way in which they were translated in the context of clinical clerkships.*

Introduction

In any medical curriculum, clinical rotations in hospitals provide essential learning situations for students to learn and practice professional clinical and cognitive skills. However, as Irby (1986) has pointed out, clinical education is flawed by problems of implementation and evaluation. First, medical teams are often so involved in patient care that they are unable to devote sufficient time and effort to the general professional education of clinical clerkships. Educating students has a relative low priority and is dominated by patient care. This seems obvious, because patient care is the primary mission of hospitals. Second, learning situations during clerkships are dependent of many accidental opportunities. Numerous variables such as patient mix, particular kind of hospital (academic or affiliated), and the number of urgent cases are perceived as factors that exert an uncontrollable influence on the nature of clinical rotations. Clinical rotations depend too much on features of the specific situation. Third, medical teams all too often do not clearly specify objectives for the knowledge, skills, values and attitudes that students should acquire during their clinical education.

An obvious question is whether it is possible, given the complexity of the clinical teaching setting, to find ways for assuring and improving clinical teaching. Carefully conducted evaluations of teaching are regarded as a valuable instrument to improve clinical education (Irby, 1993). Today, many medical schools use evaluative data for assuring and improving education. The majority of research on evaluation focuses on reliability and validity issues (Marsh, 1984). Relatively few studies exist about the effects of

evaluation on clinical education (Gijsselaers & Wolfhagen, 1990; Jolly, 1994; Wolfhagen *et al.*, 1994). Therefore, the development of approaches for evaluating clinical education should be initiated that are effective for improving ongoing programmes.

The purpose of the present paper is to describe five conditions that are clearly associated with the success of the usefulness of evaluation results. The five conditions include: willingness to adopt a critical attitude; willingness to analyse the existing situation; opportunity to discuss and carry on a dialogue; availability of a plan of action; and continuous collection of evaluative data. First, the context in which the conditions were translated is presented. Second, the way of reporting evaluation results is described. Subsequently, a description of the standards used to decide whether a particular aspect of a clerkship requires improvement is given. Next, the five conditions are outlined and the way in which they were translated in the context of clinical clerkships. Finally, a case-study is presented for illustrative purposes.

Context

At the Medical School of the University of Maastricht, The Netherlands, clinical education takes place in the form of clerkship teaching in different hospitals (academic and affiliated). Medical students are exposed for varying periods of time to the specialties of medicine: obstetrics/gynaecology, paediatrics, psychiatry, surgery, internal medicine, family medicine, and so on. Students move from one clerkship to another. Faculty members and residents are the primary teachers of clerkship students. Each clerkship teaching programme is coordinated by a clerkship coordinator. This coordinator is chairman of a team of medical teachers representing the hospitals providing a particular clerkship. This teaching team, a so-called planning group, is responsible for the organization and content of the clerkship programme in a certain discipline. These teams are fully accountable to the curriculum committee of the medical school. Evaluative data provide them with information to improve and control the quality of the clerkships.

Correspondence: H. A. P. Wolfhagen PhD, University of Maastricht, Department of Educational Development and Educational Research, PO Box 616, 6200 MD Maastricht, The Netherlands. Fax: 0-11-43-3670444, Email Wolfhagen@Educ.Unimaas.NL

Evaluation instrument

A Clinical Teaching Evaluation Questionnaire (CTEQ) was developed exclusively to evaluate clinical clerkships. It consists of a questionnaire to be filled out by students at the end of their clerkship period. For all clerkships, the same questionnaire was used, since this enables the evaluator to make comparisons between clerkships. The items of the questionnaire were derived from descriptions of how good clinical teaching can be defined. The educational conditions to be met by the clerkships were delineated by using studies aimed at characterizing clerkships and apprenticeships in vocational training and medical education (Stritter *et al.*, 1975; Irby, 1986). A final definition of required educational conditions was established by transforming the generic literature findings to the situation of the clinical clerkship. This results in 14 domains related with an effective learning environment. Each domain reflects a particular aspect of clinical education. The following domains were identified:

- (1) prior knowledge of students with regard to knowledge and skills;
- (2) received information on content and structure of the clerkship in a formal clerkship manual;
- (3) personal supervision in general;
- (4) supervision of technical performance;
- (5) attention to non-patient related affairs (context of patient situations);
- (6) quality of educational activities directly organized for student learning;
- (7) patient mix in terms of problem variation and frequency;
- (8) quality of patient-related learning situations, such as ward attendance and participation in patient rounds;
- (9) quality of outpatients' clinic regarding available office facilities for examining patients;
- (10) quality of learning facilities available, such as office facilities and learning resources;
- (11) quality of clinical competence assessment given to students at the end of the clerkship;
- (12) opportunity for self-study as a result of patient contacts;
- (13) global judgement of the clerkship quality in relation to atmosphere, organization and instructiveness; and
- (14) spread of activities over time.

In a pilot study, a preliminary version of the standard questionnaire was designed based on the domains described above. Next, all clerkship coordinators assisted in determining the content validity of the CTEQ, by responding to a draft version. Each coordinator was asked to examine the list of items (arranged around the 14 sections mentioned above) in detail. Then, pilot studies were conducted, for a period of 1 year, in two clinical departments to test content-validity and utility. Finally, the items were reorganized into a final version of the CTEQ. The final version of CTEQ contained 51 items covering 14 different domains. The majority of items were in five-point Likert type scale, ranging from (1) totally disagree to (5) totally agree. Some items required a write-in response. The CTEQ is administered to the students at the end of each clerkship.

The reliability of the questionnaire was investigated using generalizability studies, particularly investigating agreement among students (Brennan & Kane, 1979). The findings of these studies indicated that the CTEQ generally provide reliable information (Wolfhagen, 1993). The construct-validity was tested on empirical evidence through confirmatory factor analyses (LISREL). The results indicated that the model showed a reasonable fit (Chi-square (979 d.f) = 4420.06, $p = 0.000$, a root mean square residual of 0.059, a goodness-of-fit-index of 0.721 and an adjusted goodness-of-fit-index of 0.678) (Wolfhagen, 1993). Thus, the model fitted the data reasonably well.

Reporting of results

Every 6 months evaluation results are reported to clinical teachers and the Educational Committee of the Medical School. For every evaluation report mean scores are calculated across all student responses for every item. Next, these scores are aggregated to domain scores on a scale from a minimum value of 1 to a maximum value of 5. Domain scores are calculated at the level of the clerkship in a specific hospital. These domain scores were used to make decisions regarding the administration and improvement of clerkships.

Standards

The central issue in making decisions about educational quality of clerkships is whether a particular aspect of a clerkship requires improvement. A combination of relative standards and absolute standards were used to provide insights in relative strengths and weaknesses of the programme. For the relative standards two kinds of references are used. First, comparing clerkships across hospitals (e.g. surgery in hospital A versus surgery in hospital B) and second, comparing across domains within a clerkship (e.g. domain 3 versus domain 4 of surgery in hospital B). Absolute standards were used to guarantee a minimum requirement. A cut-off score of 3 (the distinction between sufficient and insufficient on a five-point scale) was chosen as an absolute standard. A particular clerkship or one of its components needs special attention if the score is significantly below the average of the scores of the reference group or below the cut-off score. On the basis of the evaluation results appropriate corrective actions were taken.

Five conditions to enhance use of evaluative data

The success of the usefulness of evaluation results is clearly associated with the fulfilment of a number of conditions. Cohen (1980), for example, showed that consultancy is a key-variable for the actual utilization of evaluation findings. Dalin (1978) mentions five conditions that need to be satisfied in a fixed sequence.

- (1) *Expressing willingness.* First of all, physicians should express their willingness to adopt a critical attitude towards their own clerkship, for example, by declaring that they want to spend time and energy on the

evaluation of the clerkships or want to attend discussion sessions.

- (2) *Analysing the situation.* Subsequently, the physicians involved should be willing to analyse the existing clerkships. This implies comparing the actual situation with the desired situation.
- (3) *Carrying on a dialogue.* All parties concerned should gain an understanding of the findings and there must be an opportunity to discuss possible solutions. Recommendations must be judged on their merits and demerits and the relationship between these two.
- (4) *Drawing up a plan of action.* The fulfilment of the preceding condition should result in a plan of action. This should include a recommendation of changes and preferably provide concrete activities. It is important that all parties involved have the opportunity to express their opinions and are willing to conform to the group's decisions. To be able to verify whether everyone has complied with them, those agreements must be committed to paper.
- (5) *Regular assessment.* The ultimate goal is that actions result in improvement. Continuous collection of evaluation data to evaluate the results of actions is, therefore, essential. Depending on the effects, further steps may need to be taken.

At the Medical School of the University of Maastricht, The Netherlands these five conditions were translated to the context of clinical clerkships. First, the Educational Committee decided that the evaluation results would be discussed during the meetings of the planning groups. As described before, a planning group is a team of medical teachers responsible for the quality of the clerkships in a certain discipline. Before implementing this procedure, a meeting was arranged between the Educational Committee and the members of the planning groups to discuss the usefulness of evaluation results and the task of planning groups in this process. All planning groups endorsed the importance of evaluation and promised to carefully consider the results. This might be seen as condition (1) Expressing willingness.

Once or twice a year the members of the evaluation project generate a report that contains the evaluation results of each clerkship. Each report provides a survey of the strong and weak aspects of a clerkship, as mentioned before. An example of a good aspect is: in Hospital C the students received very good feedback. An example of a weak aspect is: in Hospital A the supervision is insufficient; in Hospital B students do not see enough patients. A clear report might be seen as condition (2) Analysing the situation.

Each planning group received this report and discussed the evaluation results and any problems that may have been related to it. Written feedback was always accompanied by personal consultation with one of the members of the evaluation project. This type of consultation is considered important for the implementation of the evaluation findings. Research has shown that student ratings result in educational improvement only when feedback is combined with expert consultation. This might be seen as condition (3) Carrying on a dialogue.

At the end of each session, recommendations were

made by the planning group regarding the aspects of a clerkship that required change. For instance, Hospital A: actions were taken to insure that students have the opportunity to see more patients. This might be seen as condition (4) Drawing up a plan of action.

The evaluation report and recommendations of the planning group were sent to the Educational Committee. The Educational Committee has the responsibility to determine whether the actions are to be implemented and whether these actions produce the desired effect. If necessary, the Educational Committee will contact the specific hospitals. This might be seen as condition (5) Regular assessment.

In the next section a case is presented to illustrate how the fulfilment of these five conditions contributed towards the success of the usefulness of evaluation results in a paediatrics clerkship in an affiliated hospital.

The case of the paediatrics clerkship

The paediatrics clerkship takes 6 weeks during the final year of the medical programme. This case concerns a paediatrics clerkship in an affiliated hospital. During each rotation three students were doing their clerkship in this hospital. Students complained about the quality of this clerkship, but the real problem was not clear. During the meetings of the planning group the low scores for some domains of this clerkship were mentioned several times and the paediatrician promised to improve the situation; but without effect. Subsequently, the Educational Committee arranged a meeting with the paediatricians of this hospital to discuss the quality of this clerkship and to discuss how to improve its quality. Evaluation data were used to provide more detailed information about the quality of this specific clerkship.

In total, 96 (78%) students completed the evaluation questionnaire for their paediatrics clerkship; 20 (80%) referred to this particular hospital. A mean score was calculated for the different domains for this clerkship within this hospital and for the other hospitals. These scores were depicted in a profile, as shown in Figure 1. The profile was first of all used to determine the relative position of this clerkship in relation to the other hospitals. This analysis revealed that compared to the mean score of all other clerkships, this clerkship scored relatively low. Next, the profile of this clerkship was used to detect strong and weak aspects of this clerkship. As can be seen in Figure 1, for domains 4 (supervision technical performance), 5 (attention non-patient related affairs), 6 (educational activities) and 9 (outpatients' clinic), scores were relatively low and also below the cut-off score (3 on a scale 1–5).

During a meeting of the Educational Committee the problems of this clerkship were discussed in detail. Two main problems were identified: the lack of supervision and insufficient new patients available for students. It seemed that none of the paediatricians felt responsible for the students, because no supervisory appointments had been made. As a consequence, the level of supervision was inadequate. During the meeting the paediatricians decided to clarify supervisory responsibilities. During each clerkship period three paediatricians would be responsible for

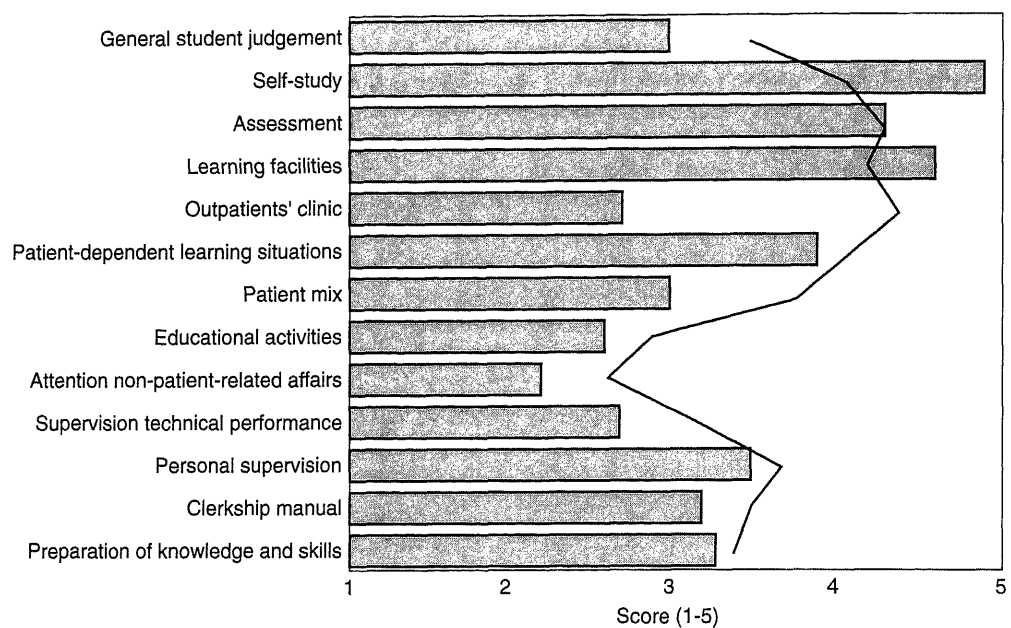


Figure 1. Profile of the paediatrics clerkship in an affiliated hospital. Average scores (scale 1–5) for the domains of the paediatrics clerkship in an affiliated hospital are depicted as bars. The average scores per domain of the other hospitals are depicted as a line.

Table 1. Average scores for the domains (scale 1–5) of the paediatrics clerkship in an affiliated hospital over four moments of measurement

Domain	Moment of measurement			
	1 N = 20	2 N = 17	3 N = 24	4 N = 18
1. Preparation knowledge and skills	3.3	3.7	3.6	3.7
2. Clerkship manual	3.2	3.5	2.9	3.5
3. Personal supervision	3.5	4.1	3.5	3.5
4. Supervision technical performance	2.7	3.3	3.1	2.6
5. Attention to non-patient related affairs	2.2	2.5	2.4	2.2
6. Educational activities	2.6	2.8	2.7	3.7
7. Patient mix	3.0	3.7	3.0	2.9
8. Patient-dependent learning situation	3.9	3.9	3.8	3.2
9. Outpatients' clinic	2.7	2.7	2.6	2.4
10. Learning facilities	4.6	4.6	4.5	4.6
11. Assessment	4.3	4.2	4.3	4.2
12. Time for self-study	4.9	4.7	4.9	4.7
13. General student judgement	3.0	3.6	3.3	3.3

the supervision, and each of the three paediatricians would supervise one student. Another problem was that students did not get the opportunity to take histories and perform physical exams on new patients at the outpatients clinic. After a discussion with the Educational Committee about this aspect, the paediatricians promised that they would give students more opportunity to see new patients.

One year later, moment of measurement two, a new profile of the clerkship was made to determine the effect of these actions. As can be seen in Table 1, the scores related to the domains 3, 4 and 5 for supervision increased as compared to moment one. As a result, they decided to continue the new organization of supervision responsi-

bilities. The score for domain 9 (outpatients' clinic) was not increased. Students got the opportunity to see more patients, but not new patients at the outpatients' clinic (during the meeting the paediatricians confirmed this). So, the effectiveness of the outpatients' clinic is still a problem.

Another year later, moment 3, the scores of most of the domains had worsened and it appeared that the supervision had deteriorated. Only the mean scores for domains 11 (assessment) and 12 (time for self-study) increased a little. Once again, the Educational Committee arranged a meeting with the paediatricians. During this meeting the paediatricians promised, not with much conviction, that they again would change the supervision. Six months later,

moment 4, a new profile was made. As can be seen in Table 1, the scores were low and there was no progress. The paediatricians then let the Educational Committee know that they wanted to terminate the contract with the University of Maastricht, because they could not agree with the educational principles of the university. They would not organize the structure of supervision to the university's wishes and they did not agree with the idea that students have to see new patients. The Educational Committee was disappointed that the quality of this clerkship did not increase, but was convinced that it was a good decision to terminate the contract with this hospital. A high-quality clerkship is important for the quality of medical education and any lessening in quality will otherwise be a waste of time and energy.

This case illustrates that evaluative data are useful to obtain detailed information about the strengths and weaknesses of a clerkship. Subjective impressions of individual students can be made more objective and results of evaluation can be used as a tool to start discussion about the basic principles of a clerkship. This case also illustrates that conditions enhancing the usefulness of evaluation results need to be fulfilled in order to be effective. During the first meeting of the Educational Committee the paediatricians expressed their willingness, condition 1 was fulfilled. The situation in the hospital was analysed, a dialogue was started and together with the Educational Committee a plan for action was made (conditions 2, 3 and 4). After a short period, the Educational Committee judged whether the actions resulted in the desired effect. In general, the quality of this clerkship improved. However, a period of progress was followed by a period of decline. Again, a meeting was organized. At that time it was concluded that the paediatricians were not really willing to reflect on changes. No discernible effects were evident. This indicates that evaluation of the clerkships needs to be a continuous process, condition 5, enabling supervisors to spot discrepancies.

Conclusion

In order to ensure that recommendations resulting from clerkship evaluation are actually used to improve the quality of clinical clerkships, five conditions need to be fulfilled. These conditions are: willingness to adopt a critical attitude; willingness to analyse the existing situation; opportunities to discuss and carry on a dialogue; availability of a plan of action; and continuous collection of evaluation data. In this article it was described how these conditions were fulfilled in a particular context. In addition, the success of the fulfilment of these conditions was illustrated by means of an example of a paediatrics clerkship in an affiliated hospital. Thus, if the conditions mentioned be-

fore are met, evaluative data can actually be used as part of a successful clerkship improvement and assurance process.

Notes on contributors

INEKE WOLFHAGEN, PhD, is an educational psychologist in the Department of Educational Development and Research for the Medical School, University of Maastricht, Maastricht, the Netherlands.

WIM GIJSELAERS, PhD, is an educational psychologist in the Department of Educational Development and Research for the Faculty of Economy, University of Maastricht, Maastricht, the Netherlands.

DIANA DOLMANS, PhD, is an educational psychologist in the Department of Educational Development and Research for the Medical School, University of Maastricht, Maastricht, the Netherlands.

GERARD ESSED, MD, PhD, is a gynaecologist in the Academic Hospital Maastricht and Professor of Clinical Education, University of Maastricht, Maastricht, the Netherlands.

HENK SCHMIDT, PhD, is a professor in the Department of Psychology, University of Maastricht, Maastricht, the Netherlands.

References

- BRENNAN, R.L. & KANE, M.T. (1979) Generalizability Theory: a review, *New Directions for Testing and Measurement*, 4, pp. 33–51.
- COHEN, P.A. (1980) Effectiveness of student-rating feedback for improving college instruction: a meta-analysis, *Research in Higher Education*, 13, pp. 321–341.
- DALIN, P. (1978) *Limits to Educational Change* (London, Macmillan).
- GIJSELAERS, W.H. & WOLFHAGEN, H.A.P. (1990) Implementing an evaluation method for quality control and quality assurance in clinical education, in: W. BENDER, R.J. HIEMSTRA, A.J.J.A. SCHERPBIER & R.P. ZWIERSTRA (Eds) *Teaching and Assessing Clinical Competence* (Groningen, Boekwerk Publishers).
- IRBY, D.M. (1986) Clinical teaching and the clinical teacher, *Journal of Medical Education*, 61, pp. 35–45.
- IRBY, D.M., (1993) Faculty Development and Academic Vitality, *Academic Medicine*, 68, pp. 660–763.
- JOLLY, B.C. (1994) *Bedside manners. Teaching and learning in the hospital setting*. PhD Thesis (Maastricht, the Netherlands: Universitaire pers Maastricht).
- MARSH, H.W. (1984) Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility, *Journal of Educational Psychology*, 76, pp. 707–754.
- STRITTER, F.T., HAIN, J.D. & GRIMES, D.A. (1975) Clinical teaching re-examined, *Journal of Medical Education*, 50, pp. 876–882.
- WOLFHAGEN, H.A.P. (1993) *Kwaliteit van klinisch Onderwijs* [Quality of Clinical Education]. PhD Thesis (Maastricht, the Netherlands: Univeritaire Pers Maastricht).
- WOLFHAGEN, H.A.P., VLEUTEN, VAN DER, C.P.M. & ESSED, G.G.M. (1994) Improving the quality of clinical clerkships using program evaluation results, in: A.I. ROTHMAN & R. COHEN (Eds) *Proceedings of The Sixth Ottawa Conference on Medical Education* (University of Toronto) pp. 484–487.