

MEDICAL EDUCATION

Can paediatric medical students devise a satisfactory standard of examination for their colleagues?

Steven Cunningham

Abstract

Objectives—To determine what standard paediatric medical students would set for examining their peers and how that would compare with the university standard.

Design—Single blinded computer marked examination with questionnaire.

Setting—University medical school.

Subjects—Medical students during their final paediatric attachment.

Interventions—Medical students asked to derive 10, five branch negatively marked multiple choice questions (MCQs) to a standard that would fail those without sufficient knowledge. Each 10 were then assessed by another student as to the degree of difficulty and the relevance to paediatrics. One year later student peers sat a mock MCQ examination derived from a random 40 questions (unaware that the mock MCQs had been derived by peers).

Main outcome measures—Comparison of marks obtained in mock and final MCQ examinations; student perception of the standard in the two examinations assessed by questionnaire.

Results—44 students derived 439 questions, of which 83% were considered an appropriate standard by a classmate. One year later 62 students sat the mock examination. Distribution of marks was better in the mock MCQ examination than the final MCQ examination. Students considered the mock questions to be a more appropriate standard (72% v 31%) and the topics more relevant (88% v 64%) to paediatric medical students. Questions were of a similar clarity in both examinations (73% v 78%).

Conclusions—Students in this study were able to derive an examination of a satisfactory standard for their peers. Involvement of students in deriving examination standards may give them a better appreciation of how standards should be set and maintained.

(Arch Dis Child 1999;80:573-575)

Keywords: medical student; multiple choice examination; standard setting

Examinations are a necessary part of the formal assessment of medical knowledge. Questions arise as to what should be examined, to what standard, and who should consider themselves responsible for setting an “appropriate” or minimum standard.¹ It is often difficult even for like minded examiners to agree a strict demarcation between pass and fail, particularly with a greater perceived importance to the individual that pass or fail would impart.

Who decides “standards” in medical education has become an important issue, not least because of demands that standards should not only be set, but maintained and possibly periodically demonstrated to an examining group of peers.

If senior medical staff are to set clinical and knowledge based standards for other senior medical staff, can other grades set standards for themselves? Student involvement in assessment is usually limited to a token presence on a committee. If given greater involvement, could medical students set their own standard of examination?

This study aimed to identify whether a knowledge based examination standard set by students themselves would be any less valid than the standard set by their tutors. The secondary and untested aims of the study were to improve understanding of how responsible standards are derived, and how questions are devised.

Methods

Paediatric medical students attend an eight week attachment to the department of child life and health at the University of Edinburgh during the second of their three clinical years. Forty four medical students were asked to devise 10 negatively marked multiple choice questions (MCQs), each containing five true/false branches (the question style most frequently used in Edinburgh). Students were familiar with negatively marked MCQs from previous examinations. They were given the following instructions.

(1) MCQs could be derived from any aspect of paediatrics that they considered relevant to an undergraduate. The questions did not have to be restricted to the Scottish paediatric core

Department of Child Life and Health, University of Edinburgh, 20 Sylvan Place, Edinburgh, EH9 1UW, UK
S Cunningham

Correspondence to:
Dr S Cunningham, Portex Anaesthesia and Respiratory Unit, Great Ormond Street Hospital for Children, Great Ormond St, London, WC1N 3JH, UK.
email: steve.cunningham@talk21.com

Accepted 18 January 1999

curriculum issued to them at the start of the course. Students could refer to any source for information.

(2) The MCQ examination had to differentiate between the following: (a) students who did not have sufficient core knowledge; (b) those who had sufficient core knowledge to be safe medical practitioners; and (c) those who excelled in their knowledge of paediatrics.

Student MCQs were handed in during the final week of their attachment, having had one week to devise their questions. Two days later a classmate was required to answer the 10 questions from each student. In answering each five branch question, classmates had to state whether they considered the topic an appropriate one for paediatric medical students and whether the question was too easy, just right, or too difficult.

A “mock” MCQ paper was derived by randomly selecting 40 questions from the 439 MCQs submitted. Questions considered irrelevant, too difficult, or too easy were excluded, and the next question in the list was used. Similarly, if a question on the same topic arose twice, the next available question in the list was used as a replacement. The 40 question MCQ paper was prepared in the same style as the final MB paper (also 40, five branch questions).

One year later a group of 76 paediatric medical students were invited to sit a “mock” MCQ paper under examination conditions. The paper comprised the 40 questions selected and took place five days before the students were due to sit their final MB paediatric MCQ examination. Students were not informed that the paper had been devised by a previous group of medical students. Standard computer marking sheets were used as in other university MCQ examinations. At the end of the examination students were asked to fill in a short questionnaire. Five days later, immediately following the students’ final examination, they were asked once again to fill in the same questionnaire. The questionnaire asked:

(1) Overall, do you feel this paper was; too easy, just right, too difficult?

(2) Overall, did you feel that the questions asked were clear and understandable?—yes/no.

(3) Overall, do you feel that the questions covered topics relevant to what medical students should know about paediatrics?—yes/no.

As with other “norm referenced” MCQ examinations, the marks obtained in the MCQ (maximum 200) were converted to a 0–5 grade, as in our final MB MCQ examination. In this system, grade 5 is given for marks > 1.5 SD of the mean norm, 4 is given for ≤ 1.5 SD but > 0.5 SD, 3 is given for ≤ 0.5 SD but > -0.5 SD, 2 for ≤ -0.5 SD but > -1.5 SD, 1 for ≤ -1.5 SD but > -2 SD, and 0 for ≤ 2 SD.

An assessment was made of how well students had managed to identify the three groups of their successor students (fail, pass, and honours).

Results

MCQ DEVISING GROUP

Forty four paediatric medical students each devised a maximum of 10 MCQs containing five branches in a negative marking model. Students provided answers to their own questions, which were used in marking the paper. The paper was intended as a student controlled examination, with student interpretation of questions and answers being paramount; therefore, student answers were not checked for “correctness”. Four hundred and thirty nine questions were submitted. Students devising questions also had to answer 10 MCQs; they considered 18 questions too easy (4%), 57 too difficult (13%), and 364 questions (83%) appropriate. Three hundred and sixty nine questions (84%) of those developed fell within the core curriculum. The remaining 70 questions (16%) covered topics not included in the core curriculum. Of the questions covering topics within the core curriculum, students considered 11% ($n = 41$) too difficult, compared with 28% of questions ($n = 20$) considered too difficult in those topics outside the core curriculum. Five students each considered one question as inappropriate for their paediatric examination (these involved polyhydramnios, testicular masses, tuberous sclerosis, poisoning, and Jones’ criteria for rheumatic fever).

MOCK MCQ EXAMINATION GROUP

One year later, 62 of 76 (82%) paediatric medical students accepted the invitation to attend a mock MCQ examination, five days before they were due to sit their final professional paediatric MCQ examination. Figure 1 demonstrates the distribution of marks obtained in the mock and final MCQ examinations, and shows that, in general, the spread of marks obtained was very good—that is, Gaussian—with a slightly wider spread of marks obtained in the mock MCQ examination.

The distribution of final MCQ marks obtained by students not attending the mock MCQ examination are also shown; these marks are slightly lower than average, though again

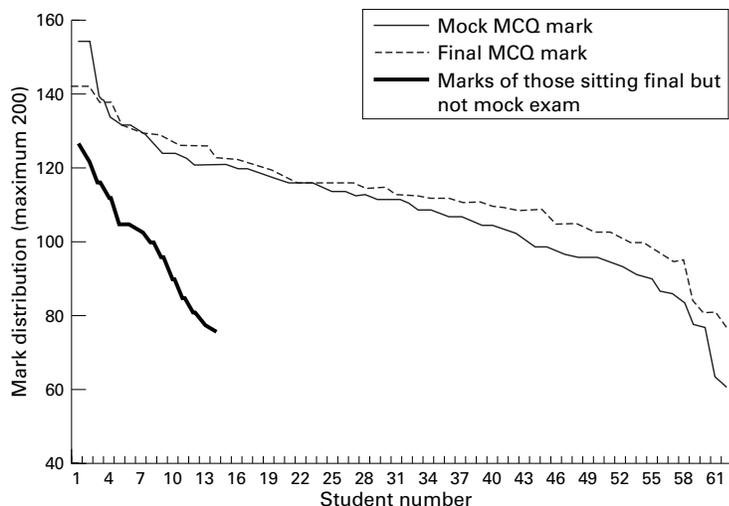


Figure 1 Distribution of marks in students taking mock and final examinations.

Table 1 Comparison of mean (SD) marks obtained by students sitting and not sitting mock MCQ examination

| | Students sitting mock MCQ | Students not sitting mock MCQ |
|---------------------------------|------------------------------|----------------------------------|
| Peripheral hospital assessment* | 3.6 (0.9)* | 3.5 (0.8)* |
| Central hospital assessment* | 3.6 (0.7)* | 3.6 (0.7)* |
| Community case assessment* | 3.3 (0.7)* | 3.1 (0.8)* |
| Long case assessment* | 3.5 (0.8)* | 3.2 (0.8)* |
| Clinical examination | 2.9 (0.9) | 2.1 (1.3) |
| Surrogate history examination | 3.2 (1.0) | 3.0 (0.7) |
| Final MCQ paper | 3.2 (1.0) | 2.1 (1.3) |

*Assessments made before mock MCQ examination.

the spread is good. Those students not attending the mock MCQ examination had lower marks in three out of four assessments made before the MCQ and scored lower in all three final examination assessments (table 1).

Correlation between the two MCQ papers (mock and final) was significant, though the relation was not particularly strong ($p < 0.001$, $r = 0.451$ for total from 200 maximum mark, or $p < 0.001$, $r = 0.542$ for norm derived mark 0–5).

Taking the two MCQ examinations (mock and final) in isolation, two students (3%) failed the mock but not the final MCQ (norm derived mark ≤ 1), and one (2%) passed the mock but not the final.

As in other centres, final examinations include a combination of assessments. In Edinburgh, the MCQ paper contributes 25% to the final MB total mark. With the mock MCQ mark rather than the final MCQ mark substituted into the total MB mark, at the upper end of the range, two students would have been elevated to honours standard (total mark $> 75\%$), and one student demoted from that standard. At the lower end, two students failed with the mock MCQ mark substituted who would have passed the final examination. No student passed with the mock MCQ mark who would have failed with the final MCQ.

The questionnaire assessing student attitudes to the two papers showed that the mock examination was considered as being aimed at a more appropriate standard of assessment than the final examination (72% *v* 31%). In addition the topics covered in the mock MCQ examination were considered more relevant to undergraduate paediatrics (88% *v* 64%). The examination questions were, however, considered to be similarly clear in both examinations (73% in mock *v* 78% for final).

Discussion

This study has shown that students are capable of setting a standard that agrees with that set by their examiners. The examination derived from student questions gave a wider distribution of marks and so resulted in a stricter examination—that is, to protect the public from unsafe practitioners it is better to fail some competent students than to pass incompetent ones (the substituted mock MCQ final mark pushed two more students to a pass/fail viva voce than the final MB mark would have done, but did not pass any that the final MB would not have done).

Multiple choice examinations, though much maligned by examiners and examinees, create

Key messages

- Medical students can be involved in the process of deriving examination standards
- Standards set by medical students may be higher than those set by their tutors
- Students may set examinations which they consider more relevant to their requirements

an opportunity to test a wide spread of knowledge in a short time. The final paediatric MCQ examination in Edinburgh is devised by several paediatric clinical staff, many of whom contribute questions to Royal College membership examinations and consequently are very experienced in creating MCQ questions. Students in this study were not instructed in the art of question creation; instead they had to rely on how they would interpret their own paediatric experience in devising a question. They did this remarkably well, in that the clarity of the questions they derived was judged almost as great as those derived by their tutors using appropriate terminology and structure. In creating this student derived MCQ examination, we also gave the students the opportunity to consolidate their learning with a mock MCQ examination and increased our bank of pretested questions for use in future final MCQ examinations.²

It is possible that those students attending the mock examination were advantaged by their attendance. However, those students choosing not to sit the mock MCQ also performed poorly in other areas of assessment (table 1), so it is unlikely that mock MCQ exposure (just five days before the final examination) would have had such a dramatic effect.

Setting a standard is always arbitrary, no matter how considered the answer may be. To keep our examinations flexible and up to date, it may be appropriate to involve those who are to be examined in the process of setting the standard. This study has shown that medical students are aware of their obligations and, contrary to the fears of some, the standard they would set for themselves may even be tougher than that set by their teachers. Medical students are capable of differentiating students who should fail from those who should pass, and can even identify those whose knowledge base defines honours status. By involving medical students in the process of examination, they would learn to appreciate standards—why they are necessary, how they are defined—and, hopefully, induce in them a sense of personal responsibility to score above the standard they and their peers would set.

Thanks to all those Edinburgh medical students who enthusiastically contributed to this project.

1 Cusimano MD. Standard setting in medical education. *Acad Med* 1996;71(suppl):S112–20.

2 McLeod PJ, Snell L. Student generated MCQs. *Medical Teacher* 1996;18:23–5.