

MRCP(UK) PART 2 CLINICAL EXAMINATION (PACES): A REVIEW OF THE FIRST FOUR DIETS (JUNE 2001–JULY 2002)

J Dacre, GM Besser and P White on behalf of the MRCP(UK) Clinical Examining Board*

INTRODUCTION

The MRCP(UK) Part 2 Clinical Examination, Practical Assessment of Clinical Examination Skills (PACES), is the new clinical component of the MRCP(UK) Part 2 Examination, which is now taken separately from the MRCP(UK) Part 2 written examination. The new-style clinical examination was introduced in June 2001, following an extensive evaluation of the existing examination by the Federation of Royal Colleges of Physicians of the UK. This paper outlines the academic rationale for the development of the new clinical examination and discusses the implementation of the examination both within and outwith the UK. The preliminary results from the first four diets (examinations sessions) are presented and proposals for further analysis are outlined.

DEVELOPMENT OF THE MRCP(UK) PART 2 CLINICAL EXAMINATION (PACES)

The new MRCP(UK) Part 2 Clinical Examination (PACES) was developed during an internal review by the MRCP(UK) Policy Committee, and after interest expressed in medical journals.^{1, 2} The review of the MRCP(UK) Examination was published in 1997.³ The intention was to make the clinical examination contemporary, evidence-based, fit for its purpose and appropriate for the examination of physicians in the twenty-first century. The areas of competence being tested in the new examination have not changed, with the exception of a greater emphasis on history taking and communication skills. Development of the new examination was informed by current educational literature and the changes in undergraduate examinations.⁴

For an examination to be robust it needs to fulfil three criteria:⁵

- validity;
- reliability; and
- practicability.

The *validity* of an examination is a term used to describe whether an examination measures what it is supposed to measure, i.e. does its content reflect the roles that need to be undertaken by those who have passed the examination (medical senior house officers (SHOs) and specialist registrars (SpRs))?

Reliability is an expression of the consistency and precision of an examination. If one considers the candidates' body of knowledge and skills, a reliable examination samples these in an appropriate and unbiased way. It is a measure of the 'signal-to-noise' ratio of the assessment. Standard clinical tests of competence, like the old long case, have an intrinsically low reliability, thought to be due to a combination of content specificity and marker variance. ('Content specificity' is a term used to explain the low correlation that exists between the ability of a candidate to perform in apparently similar components in competency-based tests; i.e. good performance in one skill does not predict good performance in a different skill. Marker variance is due to lack of standardisation of criteria against which candidates are marked and lack of training of examiners.)

Practicability recognises available resources to those who conduct assessments. With the limitations of practicability, the PACES Implementation Group set about creating a valid, reliable and practical assessment which could be implemented worldwide.

Once the outline assessment had been drawn up, pilot studies were carried out to establish the feasibility of this style of 'carousel' assessment, and to estimate the costs of its implementation on a wider scale. Five such studies were undertaken in the UK. The first two were held in January 1998 in clinical skills centres (the Whittington Hospital, London and Ninewells Hospital, Dundee), the third took place in May 1998 in an NHS out-patients department (Southern General Hospital, Glasgow), the fourth was held in March 2000 in an NHS hospital ward (Walsgrave Hospital, Coventry) and the fifth took place in March 2001 in a postgraduate centre (City Hospital, Birmingham). A total of 50 candidates, chosen from a group of SHOs working in the UK and overseas who were about to sit, or who had recently passed, the MRCP(UK) Part 2 Clinical Examination (PACES), were involved in these pilot assessments.

Pilots have also been run in every country outwith the UK that currently hosts the MRCP(UK) Part 2 Clinical Examination (PACES): Hong Kong, Kuwait, Malaysia, Oman, Singapore and the United Arab Emirates. A further pilot was held in Egypt in January 2003, before

* MRCP(UK) Clinical Examining Board: Dr Gordon Allan, Dr Eric Beck, Professor Mike Besser, Dr Nigel Bateman, Professor Jane Dacre, Dr Neil Dewhurst, Dr John Forrest, Professor Ian Gilmore, Professor Peter Kopelman, Professor James McKillop, Dr Peter Mills, Dr Robert Monie.

OCCASIONAL COMMUNICATIONS

implementing the clinical examination in May 2003.

THE PURPOSES OF THE MRCP(UK) PART 2 CLINICAL EXAMINATION (PACES)

The purposes of the MRCP(UK) Part 2 Clinical Examination (PACES) are as stated in the examination regulations.⁶ An eligible candidate must demonstrate in a clinical setting the knowledge, skills and attitudes appropriate for a physician who is completing their general professional training. The format of PACES allows a more realistic setting of day-to-day hospital practice. The published aims of the examination are outlined in Table 1.

STRUCTURE OF THE MRCP(UK) PART 2 CLINICAL EXAMINATION (PACES) CAROUSEL

The examination consists of five stations (the 'carousel'), each attended by two examiners. Candidates start at any one of the five stations and then move around the carousel of stations at 20-minute intervals (with a five-minute break between stations) until they have completed the cycle, which lasts a total of two hours.

Figure 1 shows the carousel of PACES stations. Two of the clinical stations involve the examination of two systems (Stations 1 and 3). One station (Station 5) involves the rapid assessment of four patients, similar to the previous style of short cases in the MRCP(UK) Clinical and Oral Examination. The other two stations are 'talking' stations at which the candidate is observed and marked while communicating with the patient. The mix of activities in the carousel allows an extended length of time for the assessment of communication skills, whilst preserving the most discriminating part of the old examination, the short cases.

The clinical stations (Stations 1, 3 and 5)

Stations 1, 3 and 5 test the ability of candidates to examine and discuss each of the systems in a patient. In Station 1 candidates are tested on the respiratory and abdominal systems; in Station 3, the cardiovascular and nervous systems; in Station 5, ocular disease, endocrinology, dermatology and the locomotor system. In Stations 1 and 3 candidates are allowed ten minutes

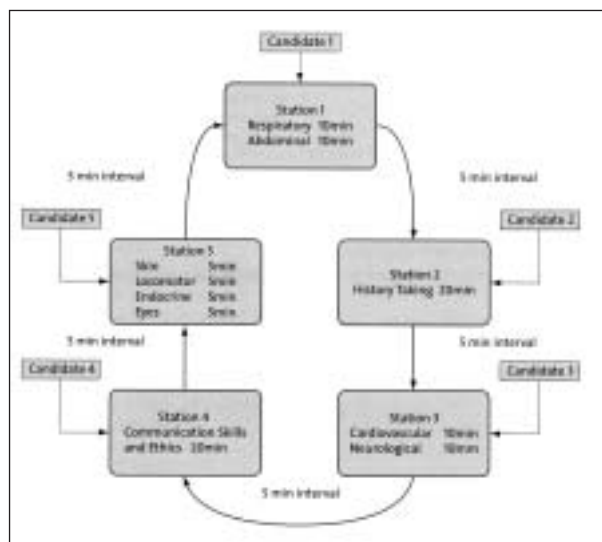


FIGURE 1
The carousel of PACES stations.

for each case and in Station 5, five minutes for each case. Before candidates assess the patient, they are given the same written instructions which introduce the patient along with a clinical question.

Examiners score each candidate against an objective marking schedule, which assesses a candidate's ability to examine a patient, to elicit and interpret physical signs and discuss the case. Figure 2 shows a copy of the respiratory mark-sheet used at Station 1. Prior to the first candidate of the cycle, the two examiners at each station examine the patients (Stations 1, 3 and 5) or talk to the patient or surrogate (Stations 2 and 4). They discuss the findings and agree what should constitute a performance warranting an award of Clear Pass, Pass, Fail or Clear Fail. This process is essential to achieve consistency of marking.

The new features of this examination are that there is a set time limit for each patient (ten minutes in Stations 1 and 3 and five minutes in Station 5), there are written instructions for each patient, and all of the systems are examined.

TABLE 1
Aims of the MRCP(UK) Part 2 Clinical Examination (PACES).

The candidates should be able to:

- demonstrate the clinical skills of history taking
- examine a patient appropriately to detect the presence or absence of physical signs
- interpret physical signs
- make appropriate diagnoses
- develop and discuss emergency, immediate and long-term management plans
- communicate clinical information to colleagues, patients or their relatives
- discuss ethical issues.

OCCASIONAL COMMUNICATIONS

The 'talking' stations (Stations 2 and 4)

Station 2 (History Taking) assesses the candidate's ability to gather data from a patient, to assimilate that information and then discuss the case. This replaces the long case in the previous Clinical and Oral Examination and extends the competencies required of successful candidates to include observed consultations.

Station 4 (Communication Skills and Ethics) assesses the candidate's ability to guide and organise an interview with the patient, discuss difficult clinical and ethical issues, provide emotional support and discuss further management. In the previous examination, this was examined in the Oral Examination.

These two stations also have detailed written instructions. In Station 2 this takes the form of a GP's letter setting out a problem in the patient to be interviewed, and for Station 4 (Communication Skills and Ethics) there is a descriptive scenario of a clinical situation, involving a communication issue or an ethical dilemma. Candidates prepare in the five minutes' waiting time between stations and are then allowed 14 minutes for the task, followed by one minute of reflection and five minutes for discussion after the patient has left the station. The two examiners are present throughout to observe the interaction.

More details of the structure of the examination are available in the introductory article about PACES published in the London College journal⁷ and in the MRCP(UK) clinical guidelines.⁸

DEVELOPMENT OF THE STATIONS

Clinical stations

Based on examiners' experience in the short cases in the old-style MRCP(UK) Part 2 Clinical Examination, standardised marking schedules were developed for each of the body systems (respiratory, abdominal, cardiovascular, central nervous system examination and skin, locomotor, endocrine and eye examination). The cardiovascular and neurological substations were paired (Station 3), as were respiratory and abdominal substations (Station 1). Those cases which had been used previously for quick diagnoses were grouped together as Station 5, with the candidate being allowed five minutes to examine one case on each system (four cases) and discuss these with the examiner. The major change from the old-style Short Cases Examination is the requirement that the physical and illustrated signs be discussed and interpreted within the time allocated. Although the clinical stations are more standardised than the previous Short Cases Examination, there is less change in the structure of the clinical stations than in the development of the new History Taking and the Communication Skills and Ethics Stations.

'Talking' stations: History Taking (Station 2) and Communication Skills and Ethics (Station 4)

A specific subgroup was set up to consider these stations. It was considered essential that the candidate was observed both gathering information (history taking) and giving information to a patient (communication skills). In addition, although ethics had become a mandatory part of the previous Oral Examination, it was felt that the practical application of ethics was important for physicians of the future. Advice was taken from communication skills and ethics experts working in UK medical schools. A curriculum was compiled to cover the topics discussed in the GMC guidance, *Good Medical Practice*,⁹ and to include consideration of the four underlying ethical principles of medicine: respect for the patient's autonomy; duty to do good and not to do harm; duty to act justly; and relevant legal aspects.

In order to develop a realistic bank of scenarios, a scenario template was drawn up for both 'talking' stations, comprising requirements for instructions to candidates, an outline script for a surrogate patient, and relevant guidance on marking for examiners. Host examiners are requested to submit scenarios to the Colleges prior to each PACES diet. These are vetted and edited by members of the MRCP(UK) Clinical Examining Board.

The three Royal Colleges currently have a cumulative total of approximately 1,500 History Taking and Communication Skills and Ethics scenarios held on file and more are being added all the time. Until now, the tendency has been for real patients to be used in the History Taking Station (Station 2) and trained surrogate patients¹⁰ in the Communication Skills and Ethics Station (Station 4).

MARKING SCHEDULES

Every candidate is examined by a total of ten examiners using a structured mark-sheet for each station or substation. The two examiners at each station mark the candidate entirely independently and without discussion, prior to submitting their mark-sheets. Discussion may then take place informally between examiner pairs and occurs as a group during the post-cycle debriefing.

Fourteen mark-sheets in total are completed for each candidate (one by each examiner at Stations 2, 4 and 5 and two by each examiner at Stations 1 and 3). Figure 2 shows a sample PACES mark-sheet.

All marks are recorded on a four-point grading system and are detailed on the PACES mark-sheets. These grades are converted to a numeric value of 1–4:

- Clear Pass 4
- Pass 3

OCCASIONAL COMMUNICATIONS

- Fail 2
- Clear Fail 1

overall mark of: Clear Fail, Fail, Pass or Clear Pass by referring to a series of published Anchor Statements (Table 2).

The marking schedules for PACES have been designed for the specific stations and substations and are of a similar pattern. Examiners are required to award an

Each marking schedule is divided into sections to help the examiner make judgements about the candidate

TABLE 2
Anchor Statements.

<i>System of examination</i>			
Clear pass Examines thoroughly and systematically.	Pass Examines systematically.	Fail Examines inadequately, either by omission or by lack of system.	Clear fail Examines badly and unsystematically.
<i>Language and communication skills</i>			
Clear pass Talks to patient in a structured but flexible manner, using intelligible language and avoiding jargon.	Pass Talks to patient in a mainly structured manner.	Fail Uses unstructured language and is unaware of communication problems with the patient.	Clear fail Talks to the patient in a completely unstructured way and uses technical jargon.
<i>Confidence and rapport</i>			
Clear pass Displays confidence, rapport and empathy.	Pass Demonstrates correct approach to the patient.	Fail May appear inappropriately confident, or unconfident/hesitant. Poor rapport with the patient.	Clear fail Causes the patient visible physical or mental distress and is oblivious to it.
<i>Clinical method</i>			
Clear pass Demonstrates correct and comprehensive clinical method and skills, eliciting the correct physical signs.	Pass Demonstrates majority of clinical skills correctly and elicits the majority of physical signs correctly.	Fail Misses important or obvious physical signs, resulting in poor or incorrect formulation of differential diagnosis.	Clear fail Misses or invents the majority of physical signs and is unable to appreciate their significance in solving clinical problems.
<i>Discussion and appreciation of patient's concerns</i>			
Clear pass Discusses clinical issues sensibly, spontaneously and with confidence, whilst able to negotiate and acknowledge areas of doubt/ignorance. Shows awareness of patient concerns.	Pass Majority of discussion sensible and correct, with no important errors of fact or interpretation.	Fail Inadequate appreciation of patient's problems and concerns. Large part of discussion incorrect through inadequate clinical skills or underlying ignorance.	Clear fail Is unaware of the patient's concerns or deals with them inappropriately. Demonstrates an inability to discuss, or most of the discussion is incorrect.
<i>Clinical thinking</i>			
Clear pass Clear, appropriate and professional.	Pass Able to solve the problem posed by the patient. Reasonable clinical thinking.	Fail Muddled clinical thinking.	Clear fail Poor grasp of clinical concepts and may be argumentative.

OCCASIONAL COMMUNICATIONS

more reliably. However, examiners are reminded to use only those statements relevant to the particular station. The overall mark (which is the one used in calculating the candidate's total score) is a global judgement of the candidate's performance, taking into account the sections of the form, and the Anchor Statements. This is similar to the key features approach described by Bordage.¹¹

STANDARD SETTING

The preliminary standard setting for the PACES Examination was based on the results of the pilots. The judgement was made based on those candidates who examiners felt should pass the examination and those who failed the examination, using the Anchor Statements. The consensus view was that 42 marks out of a possible total of 56 (75%) was likely to be necessary to demonstrate adequate competence in the core skills assessed at each Station and thus pass the examination. This is equivalent to each examiner giving a pass mark of 3 at every station or substation. However, it was recognised that this estimate was based on a small database and that might prove to be too high a mark. It was also agreed that a candidate would not pass the examination if they were awarded three Clear Fail grades by three separate examiners, regardless of the overall total achieved. Current performance in the examination, however, suggests that candidates are finding the tests more difficult than expected so far, and for each of the current diets the pass mark has been adjusted to 41, whilst still ensuring that the percentage of candidates passing the examination reflects the standard expected of the MRCP(UK) Examination.

TRAINING OF EXAMINERS

In order to implement the change, a series of examiner training activities were designed. A video describing and illustrating the new examination was produced and shown at a total of 46 training sessions run both within and outwith the UK. Examiners also had the opportunity to participate in a practice run of Stations 2 and 4 and to discuss the new examination with a member of the MRCP(UK) Clinical Examining Board. All hosts and other examiners taking part in the first diet of the examination were invited to attend these sessions. Now that PACES has been implemented, training takes place by observation during the live assessments. So far 1,157 examiners have been trained in and outwith the UK. Literature has been published by the Royal Colleges to support new examiners and also to make candidates aware of what to expect. The Royal Colleges acknowledge their responsibilities under the Race Relations (Amendment) Act 2000, the Disability Discrimination Act 1995 and Special Education Needs and Disabilities Act 2001, and monitor training activities and provide resources accordingly.

TRANSPARENCY

One of the hallmarks of the new examination is transparency. Publications include the new *MRCP(UK) Regulations and information for candidates 2002*,⁶ a curriculum of clinical examination skills (*MRCP(UK) Part 2 Clinical Examination (PACES) and clinical guidelines 2001/2*),⁸ the *Host examiner handbook* and *Examiner handbook* (both updated after each diet of PACES). The MRCP(UK) website, www.mrcpuk.org, contains and regularly updates all this information, including reports to examiners and the current MRCP(UK) Examination pass results. Each candidate also receives written confirmation of their result, including a detailed breakdown of their marks from each examiner for each station.

On the advice of the examiners, the MRCP(UK) Clinical Examining Board considers poorly performing candidates and, when it considers it appropriate, recommends individual candidates for counselling over their performance in the examination. Such counselling is conducted through the candidate's College of entry by Fellows identified by that College. Each College reports to the MRCP(UK) Clinical Examining Board on whether the candidate takes up the offer of counselling.

EVALUATION TO DATE (JUNE 2001 – SEPTEMBER 2002)

The examination went live in June 2001. By September 2002, 111 centres had run the examination in six countries over four diets and 4,215 candidates (342 outwith the UK and 3,873 in the UK) had taken the examination. The average pass rate was 47.74%. This compares to an average pass rate of 48.68% per diet for the last four diets of the old Clinical and Oral Examination.

Examiner pair agreement has been encouragingly consistent so far. In the second diet of the examination in 2002, examiner pairs reached complete agreement in 60.7% of cases and agreement to within one mark in 97.7% of cases. A difference in score of two marks was only seen in 2.2% of cases. Out of 8,204 paired results, there were only four cases (0.048%) with a three-mark discrepancy (each of these instances was reviewed by the MRCP(UK) Clinical Examining Board and discussed prior to results being released). The breakdown of these findings by station is shown in Table 3.

The overall mean mark for each station is shown in Table 4. This shows that the most difficult station is Station 3 (examination of the cardiovascular and central nervous systems). Interestingly, candidates are achieving similar marks in the History Taking and Communication Skills and Ethics Stations to the clinical stations. The pass rate was similar in centres within and outwith the UK (Hong Kong, Kuwait, Malaysia, Oman, Singapore) in the first four diets of PACES. This includes performance at the

OCCASIONAL COMMUNICATIONS

TABLE 3
Difference in examiner pair judgements by station.

Station	Difference in scores				Total
	0	1	2	3	
1. Respiratory system					
Judgements (n)	702	438	31	1	1,172
% within station	59.9%	37.4%	2.6%	0.1%	100.0%
1. Abdominal system					
Judgements (n)	690	449	32	1	1,172
% within station	58.9%	38.3%	2.7%	0.1%	100.0%
2. History taking					
Judgements (n)	733	412	26	1	1,172
% within station	62.5%	35.2%	2.2%	0.1%	100.0%
3. Cardiovascular system					
Judgements (n)	684	472	16	0	1,172
% within station	58.4%	40.3%	1.4%		100.0%
3. Nervous system					
Judgements (n)	688	452	32	0	1,172
% within station	58.7%	38.6%	2.7%		100.0%
4. Communication skills and ethics					
Judgements (n)	711	432	28	1	1,172
% within station	60.7%	36.9%	2.4%	0.1%	100.0%
5. Other systems					
Judgements (n)	773	384	15	0	1,172
% within station	66.0%	32.8%	1.3%		100.0%
Total					
Judgements (n)	4,981	3,039	180	4	8,204
% within station	60.7%	37.0%	2.2%	0.05%	100.0%

History Taking and Communication Skills and Ethics Stations.

INTERNAL AUDIT OF THE EXAMINATION

Each cycle of the examination requires ten examiners for assessment purposes. The inclusion of a large number of assessors, working independently, reduces the likelihood of discrimination or bias from individual examiners. It is not possible for one examiner, or an examiner pair, to fail a candidate outright, except in extreme circumstances where for instance a candidate puts a patient in danger or causes pain (see Regulations⁶). It has been agreed that in borderline cases the MRCP(UK) Central Office will check mark-sheets where there is more than one score of Clear Fail (1). If there is any evidence of rough handling of patients, the candidate responsible will be referred to the MRCP(UK) Clinical Examining Board. An eleventh administrative

examiner is present at each cycle to ensure that all paperwork is completed satisfactorily, to peer-review examiner behaviour, and to deputise if necessary. At the end of each centre's examination, the nominated visiting examiner and host examiner audit the documentation of the examination to ensure that there are no discrepancies. The mark-sheets and associated documentation are sent to the MRCP(UK) Central Office for scanning using a computerised optical mark reader. The examination marks are presented to, and the examination pass mark agreed by, the MRCP(UK) Clinical Examining Board before results and feedback on performance are sent to candidates. From June 2002, the results of those candidates successful in PACES have been published on the MRCP(UK) website, reducing to a minimum any delay in the notification of results.

OCCASIONAL COMMUNICATIONS

TABLE 4

Cumulative mean marks at each station (four diets).

Station	Mean marks
1. Respiratory system	2.86
1. Abdominal system	2.80
2. History taking	2.84
3. Cardiovascular system	2.70
3. Nervous system	2.74
4. Communication and ethics	2.88
5. Other systems	2.86

DISCUSSION

This paper has outlined the product of a significant review of the MRCP(UK) Examination. The new MRCP(UK) Part 2 Clinical Examination (PACES) has been implemented and has worked extremely efficiently for a period of one year (four diets). The structure and content of the assessment appears appropriate. Informal feedback from examiners and candidates indicates that the examination is fairer, more transparent and more appropriate to the needs of today's physicians in training.

There are some significant practical implications of this new assessment. Although the examination can be run using similar clinical facilities, potentially there is an increased space requirement, as a maximum of four patients are required per room for Stations 1 and 3, and individual rooms or secluded areas are required for Stations 1 and 4. In addition, each examination diet requires 11 examiners, whereas in the past only six were required.

The combination of systems examined in Stations 1 and 3 was agreed following the initial pilot and adhered to for consistency. The preliminary evaluation shown here suggests that the two most difficult systems to examine (cardiovascular and neurological) are in the same station. It may be appropriate to disaggregate these systems in the future, to explore any detrimental effect from poor performance in Station 3.

Although the process of the Communication Skills and Ethics Station (Station 4) has worked well, work is in progress to audit the content of the scenarios, to ensure an adequate mix of topics, both in terms of appropriateness and relevance of content, and the age, gender and ethnic group of the surrogates.

At present, although the new format of the examination has increased the emphasis on history taking and communications skills, these 'talking' stations are still worth a smaller proportion of the marks than the stations testing clinical examination skills. As communication is recognised as being such an important area of medicine,¹¹ there is a case for the emphasis on history taking and communication skills to be increased still further.

The need for feedback to examiners has been acknowledged. It is proposed to produce a report to be sent to each examiner giving the percentage agreement with the co-examiner and a comparison of the examiner's mean score compared to the average ('hawk/dove' index).

The Royal Colleges of Physicians have agreed to invest further resources in the analysis and evaluation of the MRCP(UK) Part 2 Clinical Examination (PACES), which will be led by the MRCP(UK) Validation, Audit and Research Group (VARG). This will include the evaluation of reliability, validity and practicability of the examination. To establish overall reliability, the generalisability of the examination is being investigated.

Analysis will focus on:

- Overall mark analysis by station;
- Relationship of indicative (sub-marks) marks to overall mark awarded for each station;
- 'Hidden' weighting of each station to overall mark;
- First station performance against overall performance;
- Effect of time of day/cycle on performance.

This article was also published in the Royal College of Physicians of London's *Clinical Medicine* 2003; **3**:452-9.

REFERENCES

- 1 Examining the Royal Colleges' examiners. *Lancet* 1990; **335**:443-5.
- 2 London D. Game, set and MRCP. *Lancet* 1994; **22**; **334(8930)**:1159.
- 3 MRCP(UK) Policy Committee. *Review of the MRCP(UK) Examination*. London: Federation of Royal Colleges of Physicians of the UK; 1997; 23-30 (internal document).
- 4 Norcini JJ. The death of the long case? *BMJ* 2002; **324**:408-9.
- 5 Newble D, Cannon R. *A handbook for medical teachers*. 2nd ed. Norwell, MA: MTP Press; 1987.
- 6 Federation of Royal Colleges of Physicians of the UK. *MRCP(UK) Regulations and information for candidates, 2002 edition*. London: Federation of RCPs of the UK; 2002.
- 7 PACES: Practical Assessment of Clinical Examination Skills. The new MRCP(UK) Clinical Examination. *JR Coll Physicians Lond* 2000; **34(1)**:57-60.
- 8 Federation of Royal Colleges of Physicians of the UK. *MRCP(UK) Part 2 Clinical Examination (PACES) and clinical guidelines (2001/2)*. London: Federation of RCPs of the UK; 2002.
- 9 General Medical Council. *Good Medical Practice*. London: GMC; 2001.
- 10 Stillman PL, Regan M, Philbin M et al. Results of a survey on the use of standardised patients to teach and evaluate clinical skills. *Acad Med* 1990; **65**:288-92.
- 11 Bordage G. *Key features approach to assessing clinical decision making*. Toronto: College of Nurses of Ontario; 1996.