

FAQ for Examiners – Changes to PACES exam

GENERAL

What is the purpose of PACES?

Why change PACES?

What are the aims of change?

When are the changes being introduced?

Does this effect the positioning of the PACES examination in training?

What doesn't change in the examination?

What differences will examiners notice?

What differences will candidates notice?

CHANGES TO THE ASSESSMENT METHOD

What changes have been made to the assessment method?

Which skills are assessed at which encounters?

THE NEW MARK SHEETS

What will be different about the new mark sheets?

Why change to a three-point marking scale?

Why is there now no "overall judgment" mark?

What has happened to the anchor statements?

How do examiners make judgements about performance in each skill?

Does calibration change?

What standard is required of the candidates?

NEW STATION 5 – INTEGRATED CLINICAL ASSESSMENT

Why change Station 5?

What is the structure of the new Station?

What are the advantages of the new format?

Will the new Station 5 be harder or easier for hosts to deliver?

Will there be any differences overseas?

How was the Station received in the pilot examinations?

Must the candidates perform history taking first?

Will it no longer be necessary for candidates to undertake examination of the optic fundus?

What happens to the dermatology, rheumatology, and endocrinology cases in the current Station 5?

Can system content already represented in other Stations be included?

Is new Station 5 easier or harder to pass than the current version?

Who will generate the scenarios?

How many skills are assessed at Station 5?

Is two minutes long enough for questioning?

DEFINING THE OVERALL PASS STANDARD – THE NEW ASSESSMENT SYSTEM

What is the current pass standard and pass rate?

How does the current marking system work?

What does “compensated marking” mean?

What is the new system?

Will this system be used immediately?

Are there any other advantages of the new system?

Are each of the seven skills of equal importance?

Is the examination more difficult or easier to pass overall?

What will happen to the pass rate?

Will examiners still know how candidates have performed overall on the day?

SPECIFIC MARKING ISSUES

How do the old and new marking scales compare?

Where is ethical knowledge assessed?

How can deficiency in medical knowledge be penalised?

How should examiners manage the patient:candidate encounter?

How should examiners manage the questioning period?

What happens if a skill is not assessed?

How do we deal with sequential errors?

Can any single examiner’s marks fail a candidate?

How is “roughness” assessed?

GENERAL

What is the purpose of PACES?

The assessment system for Core Medical Training in the UK is comprised of Workplace Based Assessment and the MRCP(UK) examination, of which PACES forms one part. The purpose of the MRCP(UK) is to allow trainees in medicine to demonstrate attainment of the knowledge, skills, behaviours and attitudes necessary for the practice of General Internal Medicine as described in the new curricula for medicine. The purpose of PACES is to assess the trainee's ability to take and interpret a clinical history, to perform a clinical examination, to apply problem solving skills to information gathered from history and examination, and to communicate sensitively and effectively with patients, carers and colleagues.

Why change PACES?

The PACES assessment format has proven to be popular, reliable and easily deliverable around the world since its introduction in 2001. The CEB wish to minimise change to the basic format, and only one of the five Stations changes in 2009.

The changes have been driven by internal quality review processes in MRCP(UK), and as a response to the changing training environment for medicine in the UK, in particular the introduction of new medical curricula, workplace assessments and PMETB standards for postgraduate assessments. The new changes improve the utility of Station 5, the method of assessment and the means by which a pass:fail standard is set, and ensure that the examination maps clearly to the new curricula, is complementary to workplace assessments and meets PMETB standards.

What are the aims of change?

There are three key aims. Firstly, to redefine the clinical skills assessed in the examination and focus examiners and candidates on the assessment and demonstration of these skills. Secondly, to relate what is assessed in PACES explicitly to the new curricula and ensure that the information regarding trainees that PACES provides is complementary to that provided by workplace assessments. Thirdly, to further enhance the fairness the method by which candidates are assessed and the pass standard set.

When are the changes being introduced?

The format of the examination changes in the third diet of 2009 (2009/3) with the introduction of the new mark sheets and the restructured Station 5. The pass:fail standard setting methodology will change in a staged manner, described in a later section.

Does this effect the positioning of the PACES examination in training?

The changes to PACES have no direct effect on the positioning of the examination in UK postgraduate training or on the rules regarding eligibility to sit, which can be found on the MRCP(UK) website.

What doesn't change in the examination?

Many elements of the examination do not change. Examiners continue to work in pairs, calibrating in detail before assessing candidates and marking each candidate independently. The content of Stations 1, 2, 3 and 4 is unaltered, as is the overall duration of the examination (125 minutes). Assessment continues to be based on observation of the candidate patient interaction, followed by a discussion of methods, findings, diagnoses and management. Importantly, the overall standard of the examination does not change – examiners should not change the overall level at which they pitch their assessment of candidates.

What differences will examiners notice?

There are three main changes. The structure and content of Station 5, the method of marking individual candidates and the means by which a pass standard is set. Examiners are required to familiarise themselves with the new format of Station 5 and the changes to the marksheets described in detail below. In particular examiners must appreciate that they must mark each skill explicitly and do not need to construct an overall judgement mark. Examiners will be aware of the change in pass:fail standard setting methodology when they meet for each post cycle debriefing.

What differences will candidates notice?

The only visible difference to candidates who have sat the examination previously is in the structure of Station 5. They will need to ensure that they are familiar with the new format and a substantial amount of explanatory material, including a new candidates' video is available for them. After the examination failing candidates will notice that feedback on their performance is provided in a different and more detailed manner.

CHANGES TO THE ASSESSMENT METHOD

What changes have been made to the assessment method?

Seven *core clinical skills*, closely related to the previous anchor statement skills, will be assessed and are shown in **Table 1**. These skills are important, easily recognisable by examiners, candidates and lay people and applicable to the wide range of clinical contexts in which physicians must practice.

Examiners assess between four and seven of these specific clinical skills at each encounter. A *three point marking scale*, rather than the current four point scale, is used, and examiners do not have to award an additional overall judgement mark for each encounter

Seven Clinical Skills in the PACES examination		
	Clinical Skill	Skill Descriptor
A	Physical Examination	Demonstrate correct, thorough, systematic, (or focused in Station 5 encounters), appropriate, fluent, and professional technique of physical examination.
B	Identifying Physical Signs	Identify physical signs correctly, and not find physical signs that are not present.
C	Clinical Communication	Elicit a clinical history relevant to the patient's complaints, in a systematic, thorough (or focused in Station 5 encounters), fluent and professional manner. Explain relevant clinical information in an accurate, clear, structured, comprehensive, fluent and professional manner.
D	Differential Diagnosis	Create a sensible differential diagnosis for a patient that the candidate has personally clinically assessed.
E	Clinical Judgement	Select or negotiate a sensible and appropriate management plan for a patient, relative or clinical situation. Select appropriate investigations or treatments for a patient that the candidate has personally clinically assessed. Apply clinical knowledge, including knowledge of law and ethics, to the case.
F	Managing Patients' Concerns	Seek, detect, acknowledge and address patients' or relatives' concerns. Listen to a patient or relative, confirm their understanding of the

		matter under discussion and demonstrate empathy.
G	Maintaining Patient Welfare	Treat a patient or relative respectfully and sensitively and in a manner that ensures their comfort, safety and dignity.

Which skills are assessed at which encounters?

Table 2 shows the skills assessed at each encounter. Each skill is assessed between 8 and 16 times for each candidate over the PACES carousel.

Station	Encounter	Skills assessed
1	Respiratory	A:B:D:E:G
1	Abdomen	A:B:D:E:G
2	History	C:D:E:F:G
3	Cardiovascular	A:B:D:E:G
3	Nervous system	A:B:D:E:G
4	Communication	C:E:F:G
5	New Station 5 (1)	All seven
5	New Station 5 (2)	All seven

THE NEW MARK SHEETS

What will be different about the new mark sheets?

The overall appearance of the new mark sheets is very similar to the current style. Each of the skills assessed at the encounter is described on the relevant mark sheet, and each skill must be explicitly marked on a scale of Unsatisfactory (0 marks), Borderline (1 mark) and Satisfactory (2 marks). Examiners are no longer required to award an overall judgement mark for the encounter.

Space for comments will still be available.

New features include colour flash encounter coding, a box to indicate which examiner took the lead in questioning, and a box to indicate which scenario was used at Stations 2, 4 and 5.

Why change to a three-point marking scale?

PACES requires candidates to demonstrate a series of clinical skills. Examiners are asked to decide whether performance in any given skill is satisfactory or not. We do not believe that further grading of a satisfactory or unsatisfactory performance in any skill will further inform overall decision-making regarding each candidate. However, in recognition of the fact that examiners may feel that a candidate did not do enough to warrant a satisfactory award for a particular skill, but should get some credit for what they did do, we have included a borderline grade. In the pilot examinations examiners used this grade in under 10% of judgements.

Why is there now no “overall judgment” mark?

We want examiners to focus on the assessment of each of the clinical skills, and not on the construction of an overall mark, in which they may weigh the relative importance of different skills demonstrated in the encounter differently. Marking skills explicitly and accurately also allows candidates to see how their performance in different skills varied.

What has happened to the anchor statements?

The anchor statements form the basis of the seven core clinical skills and are now integrated into the mark sheets.

How do examiners make judgements about performance in each skill?

Examiners base their judgements on the performance descriptors for each skill that are found on each mark sheet and their own calibration records.

Does calibration change?

Calibration is the standard setting process for the PACES examination and remains a critical part of the assessment process. Examiners must see and assess all the cases themselves before assessing the candidates and document the performance criteria needed to attain a satisfactory award in *each skill*. New calibration sheets exist for all encounters and examiners should use these to structure the discussion with their co-examiner and their subsequent assessment of each candidate.

What standard is required of the candidates?

Examiners should pitch the level of their assessment of each skill at the standard required of a trainee who is ready to exit from core medical training and enter higher specialty training. Examiners may find it useful to ask themselves “Is this candidate demonstrating this skill at the level required to be Receiving Medical Officer for the Acute Medical take?”

NEW STATION 5 – INTEGRATED CLINICAL ASSESSMENT

Why change Station 5?

The current Station 5 is less discriminating than other Stations. Although four systems are represented the cases are frequently highly predictable. Assessment focuses on diagnosis, without recourse to history, and subsequent discussion with examiners is less consistent in level and content than at other Stations. Some centres find it difficult to find enough cases, and others find it difficult to accommodate patients in a manner that provides the appropriate degree of privacy and dignity.

In addition, change to the new model permits further assessment of communication skills.

What is the structure of the new Station?

The new Station 5 is known as “Integrated Clinical Assessment” and is comprised of two ten-minute encounters, each known as a “Brief Clinical Consultation.” The candidate has a brief introductory explanatory referral (similar to Stations 2 and 4), delineating a specific problem to address, and then has 8 minutes to undertake whatever history and examination they believe are most likely to solve the clinical problem posed, answer any questions the patient may have and explain their investigation and/or treatment plan to the patient. The remaining 2 minutes is spent with the examiners, and allows the candidate to relate the relevant physical findings and differential diagnosis – information they may not have related in detail to the patient.

What are the advantages of the new format?

The new format more accurately reflects what trainees do in real practice, and for the first time in the examination candidates have to demonstrate integrated history taking and examination skills. Although the majority of encounters will still relate to one of the four current systems represented at Station 5, other content, for example relating to acute medicine, haematology, infectious disease or elderly medicine, can now be included.

Will the new Station 5 be harder or easier for hosts to deliver?

The overall patient requirements for a PACES cycle will decrease, making it easier to find and accommodate patients for the examination. Surrogates without physical signs can also be utilised at Station 5, although not exclusively. Hosts will have to generate the short scenarios for each encounter but examples will be available from examination offices.

Will there be any differences overseas?

Overseas centres that have difficulty in recruiting English speaking patients with physical signs will have the additional option of using an English speaking surrogate along with the non-English speaking patient. In this situation the candidate will examine the patient but ask the surrogate for any historical information – analogous to the real life situation when a doctor assesses a patient with communication difficulty of any sort.

How was the Station received in the pilot examinations?

Both candidates and examiners felt the new Station 5 was an improvement on the current format. It was felt that the encounters were realistic, varied and intuitive for trainees to undertake. Clear instructions regarding the precise task to be undertaken must be given in the scenario and both candidates and examiners must appreciate that, in contrast to Stations 1,2 and 3, a comprehensive and systematic examination and history is not required or possible in the time available – the aim is to assess candidates ability to perform targeted assessment.

Must the candidates perform history taking first?

No, they can use the time available any way they wish. Some may choose to examine the relevant area first, then ask questions, or vice versa. Some may wish to do what they often do in real life, and take a history and examine the patient concurrently.

Will it no longer be necessary for candidates to undertake examination of the optic fundus?

Fundal examination will no longer be a component of every PACES cycle. However, candidates will still have to be prepared to undertake it, as the focus of a Brief Clinical Consultation could be a vision related symptom.

What happens to the dermatology, rheumatology, and endocrinology cases in the current Station 5?

Again, although it will no longer be mandatory for a candidate to see a patient with a problem focussed on these systems, they will still have to be prepared to do so, and it is likely that the majority of Station 5 encounters will still have a focus in these systems. The opportunity to include cases from other systems or cases that cross system boundaries outweighs the fact that these systems will not always be represented.

Can system content already represented in other Stations be included?

Yes. For example the focus of an S5 encounter could be a respiratory problem. However, no system can form the system focus of an encounter, including the Station 2 and 4 encounters, on more than two occasions in the cycle. In addition, a candidate cannot be expected to undertake a comprehensive system based examination and take a history in the time available and the problem posed must therefore limit candidate examination to a manageable portion of any major system examination.

Is new Station 5 easier or harder to pass than the current version?

The systems content of Station 5 will extend, and the assessment is more detailed, broader and more consistent than is the case than in the current format. The total marks available from these encounters increases, reflecting the number of skills to be demonstrated. Each examiner pair will ensure that the standard set at Station 5 is consistent with the overall PACES standard. Considered and fair calibration will ensure that this occurs.

Who will generate the scenarios?

Host examiners will be responsible for the generation of the Station 5 scenarios. A scenario template and example scenarios covering the “Top 20 “ symptom presentations, and the systems currently represented at Station 5 are available. Each College will provide support in the generation and vetting of scenarios tailored to the patients available at each centre.

How many skills are assessed at Station 5?

The Station 5 encounters are the only encounters where all seven skills are assessed. As such it will be the most challenging Station for examiners using the new marking system for the first time. Experience in the pilots showed that once examiners were familiar with the marksheets and the basic structure of the Station, and calibrated on a skills basis using the new calibration sheets, they were able to assess all skills in the time available.

Is two minutes long enough for questioning?

Yes. Only two skills need be assessed in the questioning period; differential diagnosis (D) and identifying physical signs (B). All other skills can be assessed during the 8 minute candidate: patient interaction. Sufficient time will also exist to clarify uncertainties in the management plan, and on any rare occasion when the patient forgets to ask the candidate their own question, request that they then do so.

DEFINING THE OVERALL PASS STANDARD – THE NEW ASSESSMENT SYSTEM

What is the current pass standard and pass rate?

The current pass mark is 41/56. As the minimum score is 14, this is more accurately regarded as 27/42 or around 64%. This pass mark has applied in all PACES diets and using this mark, the overall pass rate has varied between 40 and 50%. The pass rate for first time UK graduates is much higher, at around 70%.

How does the current marking system work?

We have called the method used in the current examination “*Compensated Encounter Based Marking.*” Each examiner pair award independent overall judgement marks on a 1-4 scale for each of seven encounters. (The four Station 5 encounters count as a single encounter for marking purposes,) These overall judgement marks are summed to create the total mark. It should be noted that although examiners are also asked to mark “sub-domains” of performance at each encounter in the current system, these marks do not contribute directly to the overall score for each candidate. Using this system, examiners are in perfect agreement on the four point scale on 60% of occasions.

What does “compensated marking” mean?

At present a candidate can pass the examination overall by making up for poor performance at one Station with better performance at another. This would be entirely acceptable if each Station or encounter was assessing the same skills, domains of performance or competencies. In PACES that is clearly not the case, as Stations 1,3, and 5 focus on physical examination and Stations 2 and 4 on talking or communication. Candidates can therefore pass the examination despite having received no pass awards from any of the four examiners at the “talking” stations.

We would like to ensure that this can no longer happen, and also be able to show that our passing candidates are demonstrably competent across the full range of clinical skills assessed. We believe that the vast majority currently are, but the “overall judgement” methodology makes it difficult to demonstrate.

What is the new system?

A prime aim of the nPACES changes has been to introduce a system in which the passing candidate can be seen to have attained the required standard in all the skills assessed. That is, a candidate cannot make up for very poor performance in a particular skill with outstanding performance in another. This requires examiners to award explicit marks for each skill assessed rather than an overall mark for each encounter, a standard then to be set for each skill, and an overall pass to be defined as the attainment of the required standard in each of the skills assessed. We have called this system “Uncompensated Skills Based Marking”.

Will this system be used immediately?

No. In order to ensure that the standard of the examination does not suddenly change at the time of introduction of the marking changes and new Station 5, a transitional phase will occur. In the first three diets of the new examination (i.e. until 2010/2), the pass standard will continue to be based on a compensatory method in which all of a candidate’s marks are summed into a total score and related to an overall pass mark. The total score is derived from the summed scores from all examiners in each skill, rather than overall judgement scores. The pass mark has a pre-defined range that relates closely to the current pass mark and has been derived by established standard setting methods. During the transitional period the performance of candidates in each of the seven core skills will be analysed in relation to overall performance and the performance of historical peer group cohorts. Minimum required standards for each skill will then be refined, enabling the introduction of the full skills based marking system in 2010/3.

Are there any other advantages of the new system?

In addition to allowing passing candidates to demonstrate that they have attained the required standards in all the necessary skills, the new system provides clear performance feedback to those who fail. Candidates who fail can easily see where their performance in certain skills was suboptimal, in addition to seeing the encounters or Stations where they underperformed. We believe such enhanced feedback will help failing candidates prepare for future attempts.

Are each of the seven skills of equal importance?

Yes, in that all will need to be passed to attain a pass in the examination overall. However the minimum required standard for the award of a pass can be varied between the skills, thus according higher or lower importance to any. For example, in the pilots we set a much higher standard (28/32) for the skill “Maintaining Patient Welfare” than the other skills. In the first three diets of the examination the Clinical Examining Board will review the performance of all candidates scoring 28 or less for this skill and may award an overall fail even if the total test score exceeds the overall pass mark.

Is the examination more difficult or easier to pass overall?

We believe the examination will be neither harder nor easier to pass. Clinical content will not change dramatically and examiners will pitch their assessment at the same level as previously.

What will happen to the pass rate?

There is no intention to change the overall standard of the PACES examination. Intuitively, one might expect the introduction of multiple pass hurdles within a single examination to lower the pass rate, but neither the analysis of performance in skills in the current examination nor the experience of the pilot examinations suggest that pass rate change is likely to be significant. The new standard setting methodology and the use of a transitional phase as described above, will ensure that dramatic pass rate variation does not occur.

Will examiners still know how candidates have performed overall on the day?

We would still like examiners to have the opportunity to review overall candidate performance at the end of each cycle. The cumulative mark sheets are more detailed than currently, and include 86 rather than 14 separate marks contributing to the candidate's total mark. We have therefore produced an electronic Excel mark sheet that facilitates mark processing on the day and will allow examiners to see how each candidate has performed. Some centres may choose to continue using a paper system.

SPECIFIC MARKING ISSUES

How do the old and new marking scales compare?

We would like examiners to consider a Satisfactory judgement in the new scale as equivalent to a Clear Pass or Pass in the old scale, and an Unsatisfactory judgement in the new scale as equivalent to a Fail or Clear Fail in the old scale. Borderline is a new and narrow grade, awarded only when examiners feel that a candidate has not done enough to deserve a Satisfactory award, but deserves some credit for what they have done. In the pilots, it was used in under 10% of judgements.

Where is ethical knowledge assessed?

Ethical knowledge is one component of clinical knowledge, and a candidate's ability to apply clinical knowledge to a particular case is assessed in Skill E – Clinical Judgement. Skill F – Managing Patient Concerns – may also require the candidate to apply ethical principles if a satisfactory outcome is to be achieved in the clinical scenario posed. In Station 4, where ethical knowledge has traditionally been assessed, examiners must ensure that the focus of the discussion with the candidate is on knowledge and understanding of communication strategies and relevant ethical issues, and NOT on the assessment of the standard care of the condition that forms the focus of the case.

How can deficiency in medical knowledge be penalised?

PACES is an assessment of core clinical skills. One of the skills assessed is Clinical Judgement (E), defined as the ability to apply knowledge to a problem, situation or case. If an examiner believes a candidate is deficient in knowledge of relevance to the problem, situation or case being assessed that should be reflected by the award of an Unsatisfactory judgement for Skill E (Clinical Judgement). If the deficiency in knowledge is such that patient safety or welfare might be jeopardised, the candidate can also be marked down in Skill G (Maintaining Patient Welfare).

How should examiners manage the patient:candidate encounter?

Examiners should ensure candidates understand the task they have been asked to perform. They should resist interrupting candidates unless they are performing the wrong task, are mishandling the patient or their time is up. In Stations 1 and 3, if the candidate finishes before 6 minutes have elapsed, they should be asked if they want to examine anything else, but questioning can start if they then confirm they have finished. In Stations 2,4 and 5, when history taking is a component, the full time with the patient should elapse (14 minutes at Stations 2 and 4, 8 minutes at Station 5) before discussion commences.

How should examiners manage the questioning period?

In discussion, examiners should interact with candidates in a way that gives the candidate the best possible chance to show what they know. Questions should be clear, and designed to assess one or more of the skills assessed. If a candidate makes a clearly incorrect diagnosis, examiners should bring candidates back to the correct diagnosis in order to assess the subsequent skill of clinical judgement. For example, "Imagine that this patient's hepatosplenomegaly is not caused by lymphoma, but parenchymal liver disease with portal hypertension. What would be the key investigations now?"

Only one examiner should ask the candidate questions. The second examiner may request clarification if they have not heard or understood a response, but should do this through the lead examiner. If the lead examiner “dries up”, it is acceptable for the second examiner to ask questions.

What happens if a skill is not assessed?

The onus is on the candidates to demonstrate each skill, and if they fail to do so they will receive an Unsatisfactory judgement for that skill. This is most likely to happen for the Clinical Judgement skill in Stations 1 and 3, where slow presentation of findings and discussion of differential diagnosis may leave little time for discussion of investigation and treatment and assessment of Skill E, Clinical Judgement. Examiners must therefore manage the time (minimum 4 minutes) at these encounters effectively and not continue to ask questions pertinent to only one skill (e.g. differential diagnosis) when it is clear the candidate has failed the skill.

At Station 5, if the patient fails to ask questions that enable examiners to assess Skill F, Managing Patient Concerns, it is acceptable for examiners to ask the questions during the questioning period. If they omit to do so then “Not Tested” should be entered into the comment box and the lozenges left blank. *This is the only situation when Not Tested can be entered into the comments box.*

How do we deal with sequential errors?

Sequential errors are said to occur when, for example, an incorrect diagnosis leads to the suggestion of incorrect treatment, for that patient, even though the treatment suggested for the incorrect diagnosis is in itself correct. Although a theoretical problem in a skills based marking system, it occurs infrequently in real examining practice. It can also be minimised if examiners do not persist in asking the candidate questions relevant an error they have already clearly made. Two rules should apply. Firstly, the candidate should be marked with regard to their management of the patient they have seen, and their application of knowledge and skills to the solution of the problem posed in the scenario or case introducer for this patient. Secondly, examiners should at all times bring the candidate back to the problem posed, and, if necessary, the correct diagnosis, when assessing the clinical judgement skill.

Can any single examiner’s marks fail a candidate?

No. As is currently the case, no single examiner’s judgements can result in an overall fail for the candidate. Although the “three clear fails from three separate examiners” rule will no longer exist, the marking structure (in both transitional and final forms) is such that the award of minimum marks from three examiners would invariably result in a fail. (In addition it should be noted that the current three clear fail rule results in the failure of an extremely small number of candidates.)

How is “roughness” assessed?

If a candidate acts in a way that causes a patient emotional or physical discomfort, or endangers patient safety, an examiner should award an Unsatisfactory judgement for Skill G - Maintaining Patient Welfare. As is currently the case, if an examiner is in doubt about whether actual “roughness” occurred, they may discuss this with their co-examiner and/or the patient before completing their marksheet for Skill G. If two examiners agree that roughness has occurred, they should each award Unsatisfactory for Skill G. This should be raised at the post cycle meeting and will result in the candidate scoring 28/32 for Skill G. Marks of this level or below will be automatically discussed by CEB and are likely to result in failure of the candidate irrespective of the total test score or performance in other Skills.