



Review of the GMC's PLAB test: final report

Working with doctors Working for patients

General
Medical
Council

Contents

Executive summary	01
What is the PLAB test?	08
Why is this review needed?	09
How we carried out the review	10
Our findings from the review	12
Maintaining confidence in the PLAB test	44
Annex 1: Terms of membership and reference	45
Annex 2: List of recommendations	49
Appendix: Glossary of terms	52

Executive summary

The Professional and Linguistic Assessments Board (PLAB) test is designed to test candidates' ability to practise medicine at the level expected at the end of the first year of the Foundation Programme (F1). This is when graduates of UK medical schools are granted full registration.

The General Medical Council (GMC) reviews the PLAB test from time to time to make sure that it remains fit for purpose. These reviews aim to make sure that the test continues to command the confidence of those who have a vested interest in it, such as employers, educators and trainers, and patients and the public.

They need the PLAB test to be an objective, fair, non-discriminatory and fit-for-purpose method of assessing the knowledge and skills of international medical graduates who apply for registration with the GMC.

Our terms of reference for this review fell under four broad themes:

- ensuring standards
- content
- confidence
- outcomes.

“ We are concerned that employers don’t fully understand the purpose of the test”

Ensuring standards and content

Extending the scope of the test

The PLAB test’s blueprint, which is in effect the test’s curriculum, is mapped appropriately against:

- the GMC’s core guidance for doctors, *Good medical practice*
- the outcomes expected at the end of F1 training in the GMC’s guidance, *The Trainee Doctor*
- the Foundation Programme curriculum.

However, the two-part format of the test limits the range of values and principles in *Good medical practice* that can be tested. We therefore recommend that the GMC extend the scope of the PLAB test to include an assessment of the values and principles that can’t currently be tested.

This could be achieved, for example, through a situational judgement test. But the GMC will need to explore the most appropriate way of introducing this extended assessment.

We are concerned that employers don’t fully understand the purpose of the test, what it can and can’t assess, and the level at which it is set. The GMC should therefore give greater prominence to the blueprint and the overarching statement.

It should also change the name of the test to more accurately reflect its purpose, and promote the test with employers to increase their understanding of it.

Limit test attempts and reduce the time for which passes remain current

Candidates are allowed to take Part 1 of the PLAB test as often as they need to pass, but they must pass Part 1 again if they do not pass Part 2 within four attempts.

We recommend that the GMC impose a limit of four attempts at both Part 1 and Part 2 to maintain confidence in the test. Further attempts should only be allowed if circumstances beyond the control of a candidate affect their performance, or if they first demonstrate remediation. The GMC should give candidates more-informative feedback to help them prepare better for resits, and should develop a framework for candidates to demonstrate remediation.

Some candidates will be working as doctors outside the UK while they progress through the PLAB process and register with the GMC. Some, however, will not – given the risk to patient safety of skills decay, we conclude that the GMC should reduce the timeframe for the currency of PLAB passes from three to two years.

This two-year timeframe mirrors the UK Foundation Programme Office's requirement that applicants for F1 training take a clinical assessment if they graduated more than two years previously. It also mirrors the requirement for GPs who want to join a performers' list but have been out of general practice for more than two years to take a clinical assessment.

“ The results of a generalisability study will help the GMC to decide how best to increase the reliability of Part 2”

Could the PLAB test deliver a national licensing exam?

The GMC’s Council asked us to consider the feasibility of the PLAB test delivering a national licensing examination for UK graduates towards the end of F1, in the event that they decide that a national assessment is needed at that stage of training.

We conclude that it was premature to decide whether the PLAB test is the appropriate way of delivering a future national licensing examination.

During our review, bringing forward full registration for UK graduates was proposed. Furthermore, those responsible for medical education and training in the UK have yet to agree the purpose and objectives of a national licensing examination. Both issues must be resolved before a decision on the most appropriate way of delivering a national assessment can be agreed.

Confidence

The methodologies used to set the standard and pass marks for the PLAB test are widely used in medical education and training around the world and are supported by evidence. The GMC should continue to use these methodologies, but review evidence of other methodologies used in a medical context as it emerges. The GMC should consider using item response theory and statistical equating to support the current standard setting and scoring methodologies.

Improving the reliability of the Part 2 exam

The Part 1 examination has a consistently high reliability and works well. The GMC should, however, work to improve the reliability of Part 2.

We have suggested a range of options, such as increasing the number and/or length of the objective structured clinical examination (OSCE) stations. The results of a generalisability study will help the GMC to decide how best to increase the reliability of Part 2. The GMC will, however, have to balance the competing interests of increasing reliability and the logistical feasibility of, for example, increasing the number and/or length of stations.

“ Since our review began, the GMC has introduced a formal appeal procedure”

In the interests of patient safety, the GMC should continue to apply one standard error of measurement to the Part 2 examination, given the risk to patients of false positive candidates. These are candidates who do not have the basic theoretical knowledge and practical clinical skills to practise safely but pass the test by virtue of repeated attempts either fortuitously or due to familiarity with the examination material, for example. The GMC should, however, explore options for reducing the possibility of candidates compensating between practical and examination skills stations on the one hand and communication and history taking stations on the other hand.

Reasonable adjustments and complaints are handled well

The GMC has a robust, objective and fair procedure for dealing with requests for reasonable adjustments to support candidates to take the PLAB test. The GMC should review the procedure and its guidance for candidates to make sure it reflects relevant learning from its 2012 health and disability review.

The GMC also has procedures and processes in place to deal with candidates' complaints about the PLAB test. Since our review began, the GMC has introduced a formal appeal procedure. We suggest that the GMC regularly review appeal outcomes to identify and disseminate trends, lessons learnt and development work required.

The GMC gives candidates sufficient information about the PLAB test to help them apply and prepare for it. It also directs candidates to the websites of other organisations, where they can find information about job opportunities, the cost of living in the UK and immigration requirements. We suggest that the GMC regularly review the information and guidance it provides in the light of candidates' ongoing feedback.

More female examiners should be recruited

We were concerned by the underrepresentation of women in the Part 2 examiner pool. The GMC has 135 examiners, but only 20 are women. The GMC has already done work to understand the barriers that female doctors face in becoming examiners. We recommend that the GMC try to increase the number of female examiners through the continued use of targeted recruitment campaigns.

Outcomes

Research commissioned for the review found that doctors who pass the PLAB test underperform in the membership examinations of the Royal College of Physicians (MRCP (UK)) and the Royal College of General Practitioners (MRCGP) relative to UK qualified doctors. They also underperform in the Annual Review of Competence Progression (ARCP) process.

However, the purpose of the PLAB test is not to identify whether candidates have the potential to achieve equivalent outcomes as UK graduates in postgraduate medical education and training or through medical career pathways. It is designed to test candidates' ability to practise medicine at the level expected at the end of F1 training.

“ Our recommendations should make sure that confidence in the PLAB test remains”

Why do international medical graduates underperform in these postgraduate examinations?

The reasons for the differential outcomes are not known. They might reflect the fact that international medical graduates face barriers that UK graduates don't – for example, they may take longer to find employment and tend to follow less structured career pathways. Furthermore, the issue is not unique to the UK – there is evidence of underperformance in postgraduate examinations of a similar magnitude for international medical graduates elsewhere (for example, in Canada

We were concerned by the evidence of differential outcomes. However, without an explanation for the reasons, we were unable to recommend a definitive, appropriate and proportionate course of action. We therefore recommend that the GMC and other interested parties investigate further to identify why PLAB candidates underperform in these postgraduate examinations and in the ARCP. If and when the reasons are identified, the GMC should consider any changes that are needed to the purpose and standard of the PLAB test.

Before candidates can sit the PLAB test, they must demonstrate their knowledge of English. They can do this by achieving scores in the academic version of the International English Language Testing System (IELTS) required by the GMC. The research commissioned for the review found that IELTS scores have predictive validity for ARCP outcomes. We therefore recommend that the GMC consider increasing its IELTS requirements.

Our final recommendations

Taken together, our recommendations on extending the scope of the PLAB test, improving the reliability of Part 2, limiting attempts and reducing the currency of passes in the PLAB test should make sure that confidence in the PLAB test remains – until a decision is made on whether to introduce a UK national licensing examination. We suggest that the GMC publish a progress report on taking forward our recommendations in 12–18 months.

What is the PLAB test?

Why candidates take the PLAB test

All doctors must be registered and licensed with the GMC if they want to work as doctors in the UK. How doctors are registered and licensed depends, generally, on their nationality and where they qualified. Doctors who graduate outside the UK and the EEA are known as international medical graduates.

Before the GMC will register and license them, international medical graduates must show that:

- **they have the necessary knowledge of the English language***
Most do this by achieving the GMC's required scores in the academic version of the IELTS. These are at least 7.0 in each domain (reading, writing, speaking and listening) with an overall score of at least 7.5.[†]
- **they have the knowledge and skills required for medical practice in the UK**
Most do this by passing the GMC's PLAB test.[‡]

What does the PLAB test do?

The PLAB test is designed to test the ability of international medical graduates to practise medicine at the level expected at the end of F1 training. This is when graduates of UK medical schools are granted full registration. In other words, the GMC expects PLAB candidates to show that they have the same level of theoretical medical knowledge and practical clinical skills at the same stage of training as UK medical graduates.

The test's format

There are two parts to the PLAB test. The first – known as Part 1 – is a written test of knowledge applied to the care and treatment of patients. The second – known as Part 2 – is an assessment of clinical and communication skills. Only candidates who have passed Part 1 can take Part 2.

* Section 21B(1)(d) of the *Medical Act 1983*.

† In June 2014 the GMC increased the overall score required to 7.5 from 7.0.

‡ Section 21B(1)(b) of the *Medical Act 1983*.

Why is this review needed?

The GMC reviews the PLAB test from time to time to make sure that it remains fit for purpose. The previous review took place in 2003.

This review was commissioned to make sure that the test continues to command the confidence of those who have a vested interest in it, such as employers, educators and trainers, and patients and the public. They need the PLAB test to be an objective, fair, non-discriminatory and fit-for-purpose method of assessing the knowledge and skills of international medical graduates who apply for registration with the GMC. The review's terms of reference* are at annex 1 on page 45

* In April 2012 the GMC's Council amended the terms of reference to ask us to include consideration of the feasibility of the PLAB test delivering a national licensing examination if one is introduced.

How we carried out the review

In April 2011, the GMC established an independently-led working group to undertake the review (its membership is set out in annex 1). Membership included a medical director, a doctor involved in education and training, a patient/public representative, a licensed doctor who has taken the PLAB test, and two experts to advise the working group.

The working group met nine times between November 2011 and March 2014. Our discussions were informed by a range of evidence, including:

- **a literature review***
This examined evidence on the number of times candidates are allowed to sit professional examinations (UK, Europe and elsewhere) and assessments and the periods of validity (in other words the currency) of passes in these tests. It also reviewed examination and assessment methodologies and best practice.
- **a review of the PLAB test's blueprint**
This is, effectively, the PLAB test's curriculum.
- **an analysis† by Professor Chris McManus‡ and Mr Richard Wakeford§**
This covered performance in the PLAB test and the membership examinations of the Royal College of Physicians (UK) – the MRCP (UK) – and the Royal College of General Practitioners – the MRCGP.
- **a primary research project***
This analysed performance in the PLAB test, the ARCP† and IELTS. It also examined GMC fitness to practise allegations and outcomes for these doctors. We refer to this as the Durham report.

* McLachlan JC, Illing J, Rothwell C, Margetts JK, Archer J, Shrewsbury D (2014) *Developing an evidence base for the Professional and Linguistic Assessments Board (PLAB) Test* London, GMC.

† McManus C, Wakeford R (2014) *Assessing the equivalence of PLAB graduates to UK graduates* London, GMC.

‡ Professor McManus is educational advisor to the MRCP (UK) and a member of the PLAB review working group.

§ Richard Wakeford is psychometric/assessment consultant to the MRCGP.

“ The GMC’s main statutory objective is to protect, promote and maintain the health and safety of the public”

- **an analysis of anonymised Part 2 data**
- **a GMC inhouse literature review**
This looked at the challenges faced by international medical graduates and junior doctors transitioning to the UK workplace.
- **feedback on key tasks in our terms of reference. We received this through:**
 - a call for written evidence
 - roundtable discussion groups in Northern Ireland, Scotland and Wales with people with an interest in the PLAB test, such as educators and trainers, and the British Medical Association
 - two roundtable discussion groups with members of the public (Cardiff and London)
 - a roundtable discussion group with nurses and midwives (Northern Ireland)
 - a meeting with a group of refugee doctors
 - a survey of PLAB test candidates.

We have referred to feedback relevant to our conclusions and recommendations throughout this report.

- **A presentation by an expert on the theory and development of situational judgement testing.**
- **A presentation on the use of situational judgement tests as part of the GP specialty training selection process.**

The GMC’s main statutory objective is to protect, promote and maintain the health and safety of the public. Our overriding concern has therefore been to make sure that the PLAB test continues to assess the knowledge and skills of international medical graduates to the standard necessary for practising as a doctor in the UK.

We have also kept in mind at all times our public sector equality duty under section 149 of the *Equality Act 2010*.

* Tiffin PA, Illing J, Webster LAD, McLachlan JC (2014) *The Validity of the Professional and Linguistic Assessments Board (PLAB) Exam: Research Report* London, GMC.

† The ARCP is a process that provides a formal and structured review of evidence to monitor doctors’ progress throughout each stage of medical training.

Our findings from the review

This section sets out our conclusions and recommendations under the four themes in our terms of reference. A full list of our recommendations is at annex 2 (pages 49–51). A glossary of terms used is in the appendix (pages 52–56).

“ The PLAB test is set at the level of successful completion of F1 training”

Themes 1 and 2 – Ensuring standards and content

What the PLAB test can and can't examine – the blueprint

International medical graduates who take the PLAB test have no contact with the GMC in an educational sense. Their performance in the test is a snapshot of their performance on the day they take it, rather than a measure of how well they have mastered learning objectives.

The blueprint is in effect the PLAB test's curriculum. It sets out the scope and content of the test in terms of the topics, skills and procedures that a doctor who passes the test needs to know and be able to do. It includes what the PLAB test can and can't examine.

The blueprint, as well as information for candidates on how to interpret and use it, and links to source documents and reference material are published on the GMC's website. You can find these at www.gmc-uk.org/doctors/plab/blueprint.asp.

The PLAB test is set at the level of successful completion of F1 training. The blueprint is therefore mapped against the Foundation Programme curriculum, the outcomes expected at the end of F1 training in *The Trainee Doctor*,* and the GMC's core guidance, *Good medical practice*.†

It also takes into account UK hospital episode statistics and the NHS Read Codes (so that it accurately reflects the workload of a doctor in the UK) and the Royal College of General Practitioners' training curriculum (so that testing on common, important or acute conditions and the management of long-term conditions in primary care are included‡).

Questions for each Part 1 and Part 2 examination are chosen using a sampling grid to make sure that the content is consistent.

* General Medical Council (2011) *The Trainee Doctor* London, GMC.

† General Medical Council (2013) *Good medical practice* London, GMC.

‡ The blueprint does not include the advanced duties of a general practitioner.

Good medical practice and F1 training

We approached reviewing the content of the PLAB test in two ways.

- We considered whether the blueprint is mapped appropriately against *Good medical practice* and is consistent with the outcomes at the end of F1 training set out in *The Trainee Doctor*.
- We also asked people with an interest in the test whether they think the PLAB test adequately assesses the values and principles in *Good medical practice*.

We're satisfied that the blueprint is mapped appropriately against *Good medical practice* and the Foundation Programme curriculum, that it is consistent with the outcomes expected at the end of F1 training, and that it identifies the broader professional attributes that successful candidates are expected to demonstrate in clinical practice in the UK.

We conclude that the blueprint is therefore fit for purpose. That said, we recognise that the PLAB test can't examine some of the professional and ethical attributes identified in the blueprint because of the limitations of its current two-part format.

The blueprint's overarching statement* does, however, identify these attributes. It makes it clear that although they might not be assessed during the PLAB test, these qualities are expected of doctors who attempt and pass the test, and will be appraised in the workplace.

Professional standards and ethics

Feedback we received was consistent – the PLAB test needs greater emphasis on the wider ethical values and principles that underpin good medical practice in the UK. For example:

- candidates' attitudinal attributes
- patient autonomy
- working within limits of competence and seeking advice
- working under pressure
- working in multi-disciplinary healthcare teams
- raising concerns
- probity
- how candidates deal with their own health.

Assessing these kinds of values and attributes is, however, a challenge for any medical assessment – it is not unique to the PLAB test.

* General Medical Council (2014) *GMC PLAB test blueprint* London, GMC. Available at www.gmc-uk.org/PLAB___Blueprint_overarching_statement___DC2714.pdf_48527961.pdf (accessed 20 March 2014).

“ Assessing these kinds of values and attributes is a challenge for any medical assessment”

Revalidation will make sure that all doctors working in the UK engage in a regular annual appraisal and develop and maintain their knowledge and skills throughout their career.

Also, as our review was taking place, the GMC developed the *Welcome to UK practice programme* for doctors new to UK practice. This aims to raise awareness of the ethical and professional standards expected of doctors who practise in the UK through participative events and online tools (ethical scenarios and a self-assessment tool^{*}).

We welcome the fact that the GMC has made these online resources available to PLAB candidates to help them prepare for the PLAB test. However, the GMC can't make it mandatory for candidates to participate in the *Welcome to UK practice programme* in order to book a place on the PLAB test or to join the medical register.

Given the barriers that international medical graduates face when transitioning to the UK workplace[†], it would be unfair to expect them to have the same in-depth knowledge and understanding of the professional values and principles set out in *Good medical practice* as UK graduates. However, these values and principles (which are not currently examined through the PLAB test because of the limitations of its current format) are fundamental for safe, effective and compassionate healthcare in the UK. We therefore conclude that the PLAB test does not adequately assess the full range of the professional values and principles in *Good medical practice*.

Recommendation:

The GMC should extend the scope of the PLAB test to include an assessment of the wider ethical values and principles in *Good medical practice* that the current format can't test.

* You can find the *Welcome to UK practice* web pages and tools at: www.gmc-uk.org/doctors/welcomeuklogin (accessed 20 March 2014).

† Research into these barriers is available at: www.gmc-uk.org/about/research/25054.asp (accessed 20 March 2014).

“ We welcome the work the GMC has already carried out to introduce more real-life scenarios in the Part 2 examination”

How can the GMC bridge this gap?

The literature review also identified this weakness in the PLAB test’s current format. It suggested that the GMC consider introducing a situational judgement test to bridge the gap. The Royal College of General Practitioners (RCGP), the Conference of Postgraduate Medical Deans of the United Kingdom (COPMeD) and the Committee of General Practice Education Directors (COGPED) also suggested this.

Situational judgement tests are used in the UK as part of the selection process to Foundation Programme and specialty GP training to test whether candidates have the professional attributes they need at the relevant stage of training. But some people believe that situational judgement tests only show whether candidates know what they should do, rather than what they would do if faced with the scenarios in real practice.

Does the PLAB test reflect real-life practice?

Interested parties have told us that, in their view, the PLAB test assesses the ability to recall knowledge rather than to apply it. They said it should test ability to apply knowledge in clinical settings more robustly.

We therefore welcome the work the GMC has already carried out to introduce more real-life scenarios in the Part 2 examination, such as:

- **couplet stations**
These use information identified at one station to inform action at the next station.
- **high-fidelity simulator stations**
These test candidates’ ability to deal with abnormal signs and acutely ill patients.
- **professionalism stations**
These test candidates’ ethical principles.
- **handover stations**
These test candidates’ ability to assimilate information and prioritise patients’ needs.

Recommendation:

The GMC should explore how best to extend the scope of the PLAB test to include an assessment of the professional values and principles in *Good medical practice* that aren’t currently tested. For example, this could be done by including a situational judgement test or another mechanism.

“ We were concerned that the name of the test doesn’t reflect its purpose”

Do employers understand the test’s purpose?

When we discussed the blueprint we were worried about whether employers understood what the PLAB test does and doesn’t assess, and that it is set at the level of end of F1 training.

We were also concerned that the name of the test doesn’t reflect its purpose. This could lead to misunderstanding, particularly among employers, about what the PLAB test is designed to do.

Employers are, of course, responsible for making sure that the doctors they employ have the appropriate knowledge, skills and experience for particular posts. They must not rely on a pass in the PLAB test as evidence of this.

Recommendation:

The GMC should give greater prominence to the PLAB test’s blueprint and its overarching statement to improve general understanding of the current scope of the PLAB test.

Recommendation:

The GMC should change the name of the PLAB test to reflect more accurately its purpose. For example, it could be called the GMC’s Knowledge and Clinical Skills Tests.

Recommendation:

The GMC should promote the PLAB test’s purpose with employers to increase their understanding that the test is set at the level of entry to F2 training (the second year of Foundation Programme training).

“ There was concern that candidates who need multiple attempts are less competent than those who pass first time”

The GMC should limit the number of times a candidate can attempt the test

Candidates can sit Part 1 as many times as they need to pass, provided their evidence of proficiency in English remains current. They are then allowed four attempts to pass Part 2, within three years of the date they passed Part 1.

If they don't pass Part 2 on their fourth attempt, or within three years, they must pass Part 1 again (and give further evidence of their proficiency in English, if their current evidence is more than two years old) before being allowed further attempts at Part 2.

We received consistent feedback that there should be a limit on the number of resits allowed. Opinions varied on what the limit should be, but there was concern that candidates who need multiple attempts are less competent than those who pass first time. And unlimited attempts therefore undermine confidence in the test.

Survey respondents also supported a limit – again, with different views on what the limit should be. However, there was also support to maintain the current situation, particularly among candidates who had not passed the test and those who needed more than one attempt to pass.

“ There is a risk to patient safety when doctors pass the test by virtue of repeated attempts”

What is an appropriate limit for the number of resits?

Through the literature review, we tried to find objective evidence about the number of times candidates can sit comparable professional examinations and assessments. The literature review* found that:

- the evidence base for the number of resits allowed in comparable examinations and assessments was at best weak and generally absent
- the number of attempts allowed is part of an examination's design, and time limits interact with training programmes, individual specialty requirements and different national cultures
- resitting examinations benefits some candidates, particularly candidates who only just fail
- there is moderately strong evidence to indicate that there is no further benefit after four attempts. In other words, performance tends to plateau after the fourth attempt.

The literature review also highlighted the rising chance of false positives as the number of resits increases. In other words, there is a risk to patient safety when doctors pass the test by virtue of repeated attempts, either fortuitously or due to familiarity with the examination material, and do not in fact have the basic theoretical knowledge and practical clinical skills to practise safely.

There are also costs for candidates retaking examinations, such as time and financial investment. The literature review suggested that a limit of four attempts at Part 1 and four attempts at Part 2 is a reasonable compromise between these competing interests.

* McLachlan JC, Illing J, Rothwell C, Margetts JK, Archer J, Shrewsbury D (2014) *Developing an evidence base for the Professional and Linguistic Assessments Board (PLAB) Test* London, GMC.

“ The GMC should impose a limit of four attempts at both Part 1 and Part 2 of the PLAB test”

Resit attempts relate to future fitness to practise concerns

The Durham report identified trends relevant to whether to limit resits. For example, even after controlling for age and gender, performance at the first attempt in both Part 1 and Part 2 predicted a reduced likelihood of censure for fitness to practise concerns.

The report also identified possible associations between multiple attempts at both parts of the PLAB test and the likelihood of censure for fitness to practise concerns. It found:

- a slight difference in the likelihood of censure between candidates who took Part 1 three times and those who passed on first attempt
- a large difference between candidates who passed Part 1 on their first attempt and those who had four or more attempts
- candidates who took Part 2 three times or more were far more likely to be censured than those who passed on first sitting.

There is an argument that a candidate meets the standard by achieving the required scores on the day whether at first or subsequent attempts. However, in the light of the evidence identified in the literature review and the findings in the Durham report, we believe that the GMC should impose a limit of four attempts at both Part 1 and Part 2 of the PLAB test. In our view, these limits strike the right balance between ensuring patient safety, maintaining confidence in the test and managing candidates' expectations.

A limit of four attempts is consistent and fair

Medical royal colleges impose limits on the number of attempts at their national professional examinations. In line with guidance endorsed by the GMC, candidates are allowed a maximum of six attempts but there is flexibility for individual colleges to set more stringent limits. Limiting the number of attempts at Part 1 or Part 2 of the PLAB test to four would therefore be consistent with the spirit of the guidelines on limiting attempts at UK postgraduate examinations.

We also agree, however, that there might be exceptional circumstances beyond candidates' control that affect performance on the day. In the interests of fairness, we recommend that the GMC take into account exceptional circumstances if a limit on attempts at the PLAB test is introduced.

We also conclude that, in the interests of fairness, candidates who do not pass within the limits should not automatically be precluded from taking the test again. However, we agree that the GMC should only allow a further attempt if a candidate can demonstrate that they have addressed the gaps in their performance in the PLAB test over an appropriate period. This is consistent with the GMC's guidance to the medical royal colleges where candidates have exceeded the maximum number of attempts at any component of national professional examinations.

We acknowledge that many of these unsuccessful candidates will be working or undertaking postgraduate training in healthcare systems elsewhere in the world. They might not, therefore, have access to the same level of educational support available to doctors in either professional practice or postgraduate medical education and training in the UK.

However, we agree that they should demonstrate that they have engaged in additional learning and professional development over an acceptable period of time before being allowed a further attempt at the PLAB test. This will help to improve the chances of success for candidates who haven't passed within the limits we propose and for whom there must inevitably be a question about their capability of practising safely at the level of entry to F2 training.

The literature review suggested two years as an appropriate timeframe for a period of remediation before allowing a further attempt, although there is no evidence to support this timeframe. We acknowledge the challenge for the GMC of assessing the objectivity and robustness of evidence that candidates might submit.

We also acknowledge that the opportunity to give meaningful feedback in the context of a summative examination where candidates don't have educational supervisors to help them reflect will inevitably be limited. We suggest that the GMC develop a framework within which candidates can realistically, but objectively and robustly, demonstrate sufficient remediation to allow a further attempt at Part 1 or Part 2.

Recommendation:

The GMC should impose a limit of four attempts at both Part 1 and Part 2 of the PLAB test. Further attempts should only be allowed if circumstances beyond candidates' control have affected performance or on the basis of demonstrable remediation over a period acceptable to the GMC.

“ More meaningful feedback would help candidates to identify future learning needs”

Feedback on the performance of candidates

Giving feedback to candidates on their examination or assessment results (both successful and unsuccessful) helps them to reflect on their performance and identify future learning needs. The GMC currently gives PLAB candidates their Part 1 scores. It also gives unsuccessful Part 2 candidates a breakdown of the grade they were given in each OSCE station.

However, most survey respondents and a group of refugee doctors told us that they would like better feedback on their performance in the PLAB test. This is because it would help them to prepare better for resits and to identify areas for continuing professional development in the workplace. There is therefore a desire for a greater level of feedback among candidates.

More meaningful feedback would help candidates to identify future learning needs and could help mitigate the impact of a limit on the number of attempts for candidates who need more than four attempts to pass either part of the PLAB test.

Recommendation:

The GMC should provide more meaningful feedback on performance to PLAB test candidates.

Are the current timeframes too generous?

Successful Part 1 candidates must pass Part 2 within three years of passing Part 1. If they don't, they must resit and pass Part 1 again, as well as giving further evidence of their English skills if their current evidence has expired. Generally speaking, successful Part 2 candidates must also apply for registration and a licence to practise within three years of passing Part 2.

It's not clear why some candidates don't apply for registration and a licence to practise in the UK immediately, or relatively soon, after passing Part 2, given that the PLAB test is designed to assess knowledge and skills for the purposes of registration.

Feedback we received was generally consistent – the current timeframes are too long because the knowledge and skills of some PLAB candidates could have deteriorated by the time they enter the UK workplace. There was, however, no consensus on what the timeframes should be.

Some survey respondents also supported shorter timeframes but, again, there was no consensus on appropriate timeframes. A significant minority said, however, that there should be no time restrictions. Respondents who were not registered and licensed with the GMC expressed a stronger preference for no time limit than those who were.

What is a fair timeframe?

The literature review sought to identify evidence about how long passes in comparable examinations and assessments remain current before candidates have to retake them. It found that currency is most often tied to individual examination structures and design, and can also be linked to training structures. This makes it difficult to generalise on good practice. It also found extensive evidence on skills decay, which depends on a range of variables. For example:

- skills required for complex tasks decay quicker than those required for simpler tasks
- theoretical knowledge might decay more slowly than practical skills
- the confidence and experience of the individual
- the nature of initial training and feedback
- acquiring new knowledge and relinquishing previously learned knowledge.

Some candidates will be engaged in medical practice while progressing through the PLAB process and during the time between passing Part 2 and registering with the GMC. Some, however, will not and we are particularly concerned about the risk of the currency of their knowledge and skills for patient safety.

But the issues are clearly complex – the literature review concluded that, other than the evidence on skills decay, there is a lack of significant evidence for deciding appropriate periods of currency of passes in the PLAB test.

There is, however, a precedent for assessing the clinical competence of doctors who have taken time away from medical education and professional practice in the UK. The UK Foundation Programme Office requires applicants for F1 training to take a clinical assessment if they graduated more than two years previously. And GPs who want to join a performers' list but have been out of clinical general practice for more than two years must also take a clinical assessment.

Given the risk to patient safety of skills decay, it is unrealistic for candidates to expect that passes in either part of the PLAB test will not be time-bound.

Recommendation:

The GMC should require candidates to pass Part 2 of the PLAB test within two years of passing Part 1. It should also require candidates to apply for GMC registration and a licence to practise within two years of passing Part 2.

“ It is premature to decide how to deliver a future licensing examination until its purpose and objectives have been agreed”

Could the PLAB test deliver a national licensing examination?

The prospect of using existing examinations and assessments to deliver any future national licensing examination is attractive.

Using the PLAB test would obviate the need to develop new examinations and assessments, which would have considerable development and implementation costs. But there are significant questions that need to be answered before deciding whether it would be appropriate to use the PLAB test to deliver a future national licensing examination.

The most fundamental question is: what will be the future shape of education and training for UK medical students? Proposals to bring forward the point of full registration for UK graduates are under consideration, most recently in the Shape of Training review.* If full registration is brought forward in this way, the outcomes that the GMC will require medical students to demonstrate at the end of their undergraduate education and training might change. Bringing forward full registration without a period of provisional registration might therefore have implications for the blueprint of a national licensing examination and the standard at which it is set.

It is also premature to decide how to deliver a future licensing examination until its purpose and objectives have been agreed by those responsible for medical education and training in the UK. For example, whether the examination is a summative assessment of educational achievement following an undergraduate medical programme; or whether it should assess suitability for entry to postgraduate education and training. Discussions on delivering a possible licensing examination can only begin once these principles have been agreed.

* Shape of Training (2013) *Securing the future of excellent patient care. Final report of the independent review. Led by Professor David Greenaway* London, GMC.

“ The PLAB test could, in theory at least, deliver a national licensing examination”

Would the PLAB test remain?

It is also important to note that the PLAB test will no longer exist if a national licensing examination is introduced for UK graduates. The PLAB test governance framework would not therefore be appropriate for a UK national licensing examination, although the GMC's experience of running the PLAB test could inform the development of a national licensing examination.

Medical schools, foundation schools, local education and training boards, deaneries, and employers will need to be jointly responsible, and accountable, for any future national licensing examination. The experience of professional regulators and licensing boards in other jurisdictions (for example, in the US and Canada) would also helpfully inform the development phase.

Having said that – and on the assumption that no change is made to the point when UK graduates are granted full registration – the PLAB test could, in theory at least, deliver a national licensing

examination for UK graduates towards the end of F1 training.

The blueprint is fit for purpose in terms of standard and content. It is mapped against *Good medical practice*, the Foundation Programme curriculum and the outcomes expected at the end of F1 training.

- The standard setting methods for both parts of the PLAB test are well studied in the literature and widely used to set the standards of medical examinations and assessments around the world.
- The methods used to deliver the test (multiple choice examination and OSCE) are widely used in examinations in medical education and training.

The impact of a national licensing examination on GMC resources

A detailed capacity study would be needed to help understand the pressure on existing GMC resources and to model proposals (and the likely capital and operational costs) of the GMC delivering the OSCE component of a national licensing examination.

In 2013, just over 7,500 medical students graduated from UK medical schools. It would not be difficult for the GMC to deliver a multiple choice examination to such a large cohort locally through computer-based testing.

However, an initial review of the GMC's Clinical Assessment Centre's capacity to meet the additional demand of around 7,500 UK graduates each year suggests that there would be a shortfall in capacity. Furthermore, if demand is compressed into a short period (for example, a hypothetical two-month period), logistical problems inevitably become greater. Releasing some 7,500 doctors from their posts over a short period could affect employers' ability to deliver high-quality healthcare, if a national licensing examination is taken towards the end of F1 training.

Other safety and fairness issues to consider

In addition, adequate safeguards would have to be developed to make sure that standards remain consistent and to mitigate any risk to patient safety.

- Running the OSCE component of a national licensing examination locally across the UK would introduce variables that could adversely impact standards, such as inconsistent marking by examiners across different sites, and quality of venue and equipment.
- The GMC and partner organisations would need to develop, expand and maintain the existing Part 1 and Part 2 question banks to make sure that the integrity of both examinations is not compromised. While challenging, this is not insurmountable and could be achieved, at least partly, through sharing and adapting other organisations' question banks.
- The challenges of recruiting and training sufficient OSCE examiners, and maintaining reliability, should not be underestimated.

“ The law would also need to change if full registration for UK graduates is dependent on success in a national licensing examination ”

There will also be principles of fairness and equality of opportunity to consider.

- If a national licensing examination took place towards the end of F1 training, candidates who take it too early during F1 would be disadvantaged – they will not have been given the same opportunities for learning and gaining experience compared with candidates who take it at a much later stage. This could be avoided if all candidates take the assessment on the same day, but this would affect delivery of healthcare.
- What are the implications for F1 trainees who fail? The GMC and those responsible for medical education and training would need to consider very carefully the scope for resits and an appeal procedure.

The law would also need to change if full registration for UK graduates is dependent on success in a national licensing examination. There is currently no statutory provision to require UK graduates to pass a national licensing examination before the GMC registers and licenses them.

We agree that it is too early to determine whether the PLAB test could, or should, deliver any future national licensing examination in the UK. This can only be decided once both the future shape of education and training for medical students and newly qualified doctors and the purpose and objectives of any future licensing examination are settled.

If a decision is then made that the PLAB test is the appropriate mechanism to deliver a national licensing examination, the GMC and those responsible for medical education and training will need to resolve the caveats we have identified with regard to capacity, governance, maintaining standards, and equality and fairness.

Recommendation:

The GMC and those responsible for medical education and training in the UK should agree the most appropriate way of delivering a national licensing examination when the future shape of education and training for medical students and newly qualified doctors and the purpose and objectives of the examination are settled.

Theme 3 – Confidence

Should the method of standard setting change?

The GMC currently uses the Angoff method to set the standard of the Part 1 examination, and borderline group scoring methods to set the standard of the Part 2 assessment.

The literature review concluded that 'there is no single universally optimum method of standard setting... The current methods used in PLAB are well studied in the literature, show acceptable properties, and there is no consistently better method under all circumstances'.*

The literature review highlighted other standard setting methods. It also identified other statistical approaches, such as item response theory, that can support standard setters and help develop a deeper understanding of how items and candidates perform.

Recommendation:

The GMC should retain the Angoff and borderline group scoring methods for setting the standards of the Part 1 examination and the Part 2 assessment.

Recommendation:

The GMC should consider using item response theory and statistical equating to support the current standard setting and scoring methodologies for the PLAB test.

Recommendation:

The GMC should regularly review the PLAB test's standard setting and scoring methodologies to ensure that the test remains up to date and in line with evidence of other methods in a medical context.

* McLachlan JC, Illing J, Rothwell C, Margetts JK, Archer J, Shrewsbury D (2014) *Developing an evidence base for the Professional and Linguistic Assessments Board (PLAB) Test* London, GMC (page 56).

“ The reliability of the Part 1 examination is consistently high”

Reliability of the Part 1 examination

Part 1 is a written test of medical knowledge designed to assess the application of clinical knowledge to the care of patients. It is a three-hour multiple choice examination that comprises 200 single best answer questions. It can be taken in the UK or overseas.

This examination method is very widely used in undergraduate and postgraduate medical education and training around the world. The reliability of the Part 1 examination is consistently high at 0.9 and above. (A perfectly reproducible test would have a reliability coefficient of 1.0.) The Part 1 examination therefore works well.

Computer-based support

Part 1 candidates mark their answers on answer sheets, which are then scanned using an optical mark reader. The questions in any Part 1 examination may already include the use of images, electrocardiograms and x-rays. However, the literature review noted that computer-based testing enables a wider range of assessment approaches¹ – such as the use of video or audio clips.

The literature review also noted that computer-based testing enables much quicker feedback for candidates and a greater range of data analysis.

Against these benefits, however, the literature review highlighted that it might not always be feasible to use computer-based testing if large numbers of computer stations are needed in a wide variety of locations.

Recommendation:

The GMC should explore the feasibility of introducing computer-based testing for the Part 1 examination and electronic marking for both parts of the PLAB test. This would enable the GMC to provide quicker feedback for candidates, use a wider and even more realistic range of assessment techniques, and gather additional intelligence for analysis.

* McLachlan JC, Illing J, Rothwell C, Margetts JK, Archer J, Shrewsbury D (2014) *Developing an evidence base for the Professional and Linguistic Assessments Board (PLAB) Test* London, GMC (page 56).

“ This examination format is well recognised and widely used in both undergraduate and postgraduate medical education and training”

The Part 2 examination

Part 2 is an OSCE that takes candidates through 14 five-minute stations. It is an assessment of practical, clinical and communication skills in four domains: communication, examination, history taking and practical skills. It takes place in the GMC's Clinical Assessment Centre and only candidates who have passed Part 1 can take Part 2.

This examination format is well recognised and widely used in both undergraduate and postgraduate medical education and training around the world. We support its use for the Part 2 examination. Examiners are carefully selected and well trained and are committed to the examination. The facilities at the GMC's purpose-built Clinical Assessment Centre are excellent.

Setting the pass mark for Part 2

Each station in each Part 2 examination has between three and six objectives. Each objective is weighted, with the total weightings for each station adding up to 100%. Examiners award candidates a grade between A and E for each objective but are not aware of the objectives' weightings. Grades awarded are then converted to a scale of 0–4 and calculated to give an overall station score. The examiner also gives a separate overall judgement of each candidate's overall performance (pass, borderline or fail). They base this on their professional experience using the notional 'minimally competent' doctor at the level of entry to F2 training as a benchmark.

The passing score for each station is then calculated from the mean scores of previous candidates to whom examiners have given borderline judgements. The scores for each station are added, plus one standard error of measurement (SEM), to determine the total score for each Part 2 examination.

“ Removing this margin would undoubtedly lead to candidates passing who should not”

What are the risks of false positives?

It is easy to identify good or poor candidates in any examination or assessment. The key challenge is to differentiate between the candidate who just passes and the candidate who just fails.

No assessment method will entirely remove false positives – those who pass when they should not. Many examining boards attempt to deal with this by adding a margin to the pass mark so those who might barely pass will fail.

The GMC doesn't apply any SEM to the Part 1 pass mark: candidates pass or fail based on their raw scores. However, the GMC adds one SEM to the Part 2 pass mark to minimise the chances of false positive candidates. This is imperative for patient safety, given that passing the PLAB test is a route to registration and a licence to practise, and thereby access to patients. Removing this margin would undoubtedly lead to candidates passing who should not and therefore put patients at risk of harm.

Applying any SEM may be seen as unfair to borderline candidates. However, we agree that it is nonetheless in the wider public interest. The addition of one SEM is a reasonable compromise between the risk of false positives for patients and the consequences of failure for the slightly raised number of false negatives.

Recommendation:

In the interests of patient safety, the GMC should continue to apply one SEM to the pass mark for the Part 2 examination.

“Candidates should demonstrate competence across the skill areas”

Reducing the risk of compensation

The stations in the Part 2 examination are designed to examine different skill areas: practical and examination skills and communication and history taking skills. To pass, candidates must meet or exceed the total passing score, as well as achieve the passing score in a minimum of nine (of 14) stations. This requirement was introduced to minimise the chances of candidates compensating poor performance in some stations with a very high performance in others.

However, an analysis of anonymised Part 2 data found that some candidates had compensated for poor performance in communication and history skills stations with better performance in the practical and examination skills stations. We were concerned by this, given the importance of communication skills for safe and effective patient-centred care. We agree that candidates should demonstrate competence across the skill areas and should not be able to compensate for overly-poor performance in some skill areas with better performance in others.

Recommendation:

The GMC should explore how to reduce the possibility of candidates compensating between practical and examination skills stations on the one hand and communication and history taking stations on the other hand. For example, the GMC should consider whether to require a passing standard of performance in the practical and examination skills stations and also in the communication and history taking stations.

“ Reliability is not the sole measure of an assessment”

Reliability of the Part 2 examination

The reliability of the Part 2 examination is generally around 0.65, which is typical for an OSCE of this length, but nevertheless falls below the widely accepted minimum of 0.8. This is linked to the number and length of the OSCE stations. However, reliability is not the sole measure of an assessment: other considerations such as feasibility and cost-effectiveness are also important.

Can the reliability of Part 2 be improved?

We considered ways that might increase the reliability of the Part 2 examination – for example, by increasing the number of stations in each OSCE, increasing the length of the OSCE stations or using two examiners in each station rather than one.

Feedback we received included concerns that Part 2 candidates need only pass nine of the 14 OSCE stations and that five minutes was too short a time to assess competence in a real-world clinical skill. Some said that candidates should pass all stations but, as the literature review concluded, this is unrealistic.

Increasing the number of stations in each Part 2 examination and the length of each station could increase the reliability of the Part 2 examination. The literature review also attempted to identify evidence on the optimum number of stations that should be included in the Part 2 OSCE.

However, the literature review proposed that, to help answer both these suggestions, the GMC should carry out a generalisability study, followed by a decision study (see glossary of terms in the appendix).

We also discussed whether other factors might affect the reliability of the Part 2 examination, such as:

- whether there is a sufficient link between an examiner’s mark for overall judgement and that particular candidate passing or failing
- the fact that examiners do not know the weightings given to each objective or the borderline pass marks
- the fact that examiners don’t know whether the candidates they have assessed pass or fail, or what the other examiners thought of the candidates on the day.

Reliability could be increased if a feedback session were introduced directly after each Part 2 examination. This would give examiners a chance to see the overall marks given by other examiners to all candidates, and to discuss discrepancies in the marks they have awarded.

“ Reliability could be increased if a feedback session were introduced”

However, we recognise that the benefit of introducing such an approach needs to be weighed carefully against the risk of perceived collusion among examiners. The individual objective judgement of each examiner is key to any OSCE format.

The GMC has commissioned a generalisability study. But until the results are known, we can't recommend the most appropriate approach, or combination of approaches, that the GMC should take to increase the reliability of Part 2.

Recommendation:

The GMC should seek to increase the reliability of the Part 2 examination. Options include increasing the number and/or length of the OSCE stations, introducing a feedback session after each OSCE for examiners to discuss candidates' performance, and using two examiners to assess performance. However, the GMC will need to take into account the results of the generalisability study as well as feasibility when determining the most appropriate and proportionate way of increasing the Part 2 reliability.

Equality of opportunity

As set out in the review's terms of reference, we were also asked tasked with reviewing whether the PLAB test and its administration, including complaint handling and associate recruitment, remain compliant with the *Equality Act 2010*. Given the breadth of this key task, we focused on equality of opportunity. In particular we considered:

- how the GMC deals with candidates' requests for reasonable adjustments
- how candidates can seek redress if they think something has gone wrong
- how the GMC recruits PLAB associates
- candidates' access to information about the PLAB test.

“ The GMC’s guidance to candidates explains how to request a reasonable adjustment”

How the GMC deals with requests for reasonable adjustments

The *Equality Act 2010* requires the GMC to make reasonable adjustments to either part of the PLAB test to make sure that candidates are not disadvantaged by a disability. The examination regulations include provisions for requests for reasonable adjustments as well as for review of a decision not to make a reasonable adjustment.

The GMC’s guidance to candidates explains how to request a reasonable adjustment and gives examples of adjustments that it has made for candidates. This makes sure that applicants from countries where reasonable adjustments are not routinely made know that they can ask for an adjustment.

GMC staff follow a written procedure for dealing with requests for reasonable adjustments. Where necessary, it takes into account any recommendations for an adjustment from an appropriate specialist; and gives the candidate an opportunity to agree the extent of proposed adjustments.

We treated with caution the data on reasonable adjustments made as the number of adjustments is very small when compared with the total candidate population. However, the GMC has not refused a request for a reasonable adjustment supported by an appropriate specialist; and the PLA Board chair has never had to consider a referral from a dissatisfied candidate.

The GMC’s approach to dealing with requests for reasonable adjustments is consistent with the approach agreed by the medical royal colleges’ Academy Assessment Committee and with the approach taken by most other bodies involved in medical examinations.

The GMC’s procedures for dealing with candidates’ requests for reasonable adjustments are objective, robust and fair in the context of supporting candidates with a disability to take the PLAB test.

In 2012, the GMC undertook a health and disability review, which included establishing the Health and Disability in Medical Education and Training Group. The group’s aim was to develop a comprehensive picture of the issues medical students and trainees with a disability in the UK face during their education and training. In 2013, the GMC took forward work based on the group’s findings. We welcome the GMC’s commitment to make sure that learning from its health and disability review* is reflected in its procedure and candidate guidance on reasonable adjustments.

Recommendation:

The GMC should seek to increase the number of female examiners through the continued use of targeted recruitment campaigns.

* Read more about the GMC’s health and disability reviews at: www.gmc-uk.org/healthanddisability (accessed 20 March 2014).

“ A dedicated team deals with all complaints about the PLAB test”

How do candidates raise a complaint?

Each GMC directorate deals with complaints about the services it provides, with the emphasis on addressing them promptly and effectively, and on making timely adjustments to procedures, where necessary. A dedicated team in the Registration and Revalidation directorate deals with all complaints about the PLAB test in line with a written procedure. As well as responding to individual complaints, this team analyses and reports on complaint trends to make sure that any lessons learnt are fed into service improvement projects.

The examination regulations covering both parts of the PLAB test stipulate that candidates who want to complain should do so in writing within 28 days of their examination. In practice, however, the GMC accepts complaints in alternative formats – for example, over the telephone.

The number of complaints received is very low when compared with the number of candidates who take the PLAB test. Only 293 doctors complained between 2009 and 2013. During the same period, there were just under 17,000 attempts at Part 1 and just over 9,600 attempts at Part 2.

Candidates can also ask for a clerical check of their final marks. This is an administrative check of whether marks on performance in the written examination or the practical assessment have been transcribed correctly from the completed scripts or the examiners' score sheets. As with complaints received, the number of requests for clerical checks is very low when compared with the number of candidates. The GMC received just 122 requests for clerical checks between 2009 and 2013. No wrong marks were found.

The GMC therefore has procedures in place to deal with the low number of candidates' complaints and to check that marks have been correctly transcribed. However, we were concerned that none of these procedures amounted to a fair process for candidates to appeal against their marks. In the light of our concern, the GMC has since implemented an appeal procedure.* We welcome this development.

Recommendation:

The GMC should regularly review appeal outcomes to identify and disseminate any trends, lessons learnt and development work required.

* You can find details of the appeal procedure at: www.gmc-uk.org/doctors/plab/PLABappealsreg.asp (accessed 20 March 2014).

“ We were concerned that women are underrepresented among the largest group of PLAB associates”

Women and younger doctors are underrepresented among PLAB associates

The GMC uses associates to deliver a wide range of its functions, including the PLAB test. Associates are not GMC employees but are contracted to provide services in areas of their expertise.

The demographic profiles of the PLA Board associates and the Part 1 and Part 2 Panels gave us no cause for concern. We were concerned, however, that women are underrepresented among the largest group of PLAB associates – the Part 2 examiners. There are 135 examiners, of whom the majority are men (115). There are only 20 female examiners.

Historically associate recruitment was managed locally by individual teams within the GMC. Since 2011, however, the GMC has recruited and selected all PLAB associates against defined competencies in the same way as all other GMC associates, through recruitment campaigns overseen by the GMC’s corporate HR function.

The gender balance has improved during our review through better targeting of potential female examiners, for example through raising awareness of recruitment with the Medical Women’s Federation. But the GMC recognises the need for further improvement and has already carried out work to identify the barriers faced by female doctors in becoming examiners.

This will help the GMC to understand the underlying reasons for female underrepresentation and help the GMC to take action to improve the gender balance of the examiner pool.

Younger doctors are also underrepresented in the examiner pool, although to a much lesser extent than female doctors. However, we accept that there is a balance to be struck here. Examiners must have experience of assessing performance but can only gain this as their professional experience increases. Appointing examiners who do not have sufficient assessment experience would be self-defeating and put patients at risk.

Recommendation:

The GMC should review its procedure for dealing with candidates’ requests for reasonable adjustments and its candidate guidance to make sure both reflect relevant learning from the 2012 health and disability review.

Information the GMC offers candidates before the PLAB test

The GMC provides a wealth of information about the PLAB test and advice on how to apply and prepare for it. This includes:

- what the test includes, its format and how it is marked
- the blueprint, together with information on how candidates can use and interpret it, and links to source documents
- example Part 1 questions and Part 2 OSCE scenarios
- a video on the Part 2 OSCE so that candidates know what to expect on the day
- links to the GMC's ethical guidance and interactive learning materials.

To help them consider whether life in the UK is right for them, the GMC also signposts candidates to other organisations' websites that have information about job opportunities, the cost of living in the UK and immigration requirements.

Despite this, survey respondents and a group of refugee doctors told us they would like better and more realistic information about job and training opportunities in the UK, training career pathways and other information about living and working in the UK, including the cost of living. Some also told us they would like better revision materials, including more example questions and scenarios, and a list of medical textbooks.

We recognise the significant challenges for anyone deciding to move to live and work in another country. However, it's up to individuals to research fully the implications for them of moving to the UK. Doctors are also responsible for their own learning to make sure their knowledge and skills remain up to date.

The blueprint sets out the scope of the PLAB test and is available for candidates to use as a guide to the topics on which they might be examined. And the GMC provides information for candidates on how to interpret the blueprint. As in any other summative assessment, it is up to candidates to satisfy the GMC that they have the broad range of knowledge and skills across all areas identified in the blueprint.

Recommendation:

The GMC provides sufficient information to help candidates prepare for the PLAB test and directs them appropriately to other organisations. However, the GMC should regularly review candidates' feedback to make sure that the information and guidance it provides adequately meet their needs.

Theme 4 – Outcomes

Are those who pass the PLAB test more likely to experience difficulties in education, training and practice?

We considered this through research into the performance of PLAB candidates relative to UK graduates in specialty examinations and in the ARCP, as well as data on fitness to practise cases. The findings of Professor McManus's and Mr Wakeford's study and the Durham report informed our discussion.

McManus and Wakeford* found that:

- performance in the PLAB test is a valid predictor of performance in the MRCP (UK) and the MRCGP
- PLAB candidates do not show outcome equivalence at the MRCP (UK) or the MRCGP at first attempt when compared with the median UK graduates' performance
- PLAB candidates underperform in comparison with UK graduates at the knowledge tests by approximately one standard deviation; and at the clinical assessments by approximately one and a half standard deviations.

* McManus C, Wakeford R (2014) *Assessing the equivalence of PLAB graduates to UK graduates* London, GMC.

The Durham report* found that:

- PLAB candidates performed less well at the ARCP than UK graduates
- based on an analysis of PLAB candidates' performance in IELTS,† requiring a minimum score of 8.5 in IELTS would achieve bare equivalence in the ARCP between PLAB candidates and UK graduates
- the over-representation of PLAB candidates censured in fitness to practise procedures was reduced to an insignificant level when controlled for age and gender differences
- the probability of fitness to practise outcomes is not easily predicted from PLAB scores but might be associated with the number of attempts and, to an extent, overall performance in the PLAB test.

International medical graduates have contributed over the years to the success of the NHS; and they work at all levels of practice in the UK health sector. The evidence of barriers they face on entering the UK workplace is well known.‡

The differential outcomes for PLAB candidates might reflect the fact that international medical graduates face barriers that UK graduates don't. For example, they typically take longer to find employment and tend to follow less structured career pathways.

* Tiffin PA, Illing J, Webster LAD, McLachlan JC (2014) *The validity of the Professional and Linguistic Assessments Board (PLAB) Exam: research report* London, GMC.

† Not all PLAB candidates demonstrate English language ability through achieving the scores the GMC requirements in IELTS (academic version).

‡ Slowther A, Hundt GL, Taylor R, Purkis J (2009) *Non UK qualified doctors and Good Medical Practice; The experience of working within a different professional framework*, Warwick, University of Warwick.

“ International medical graduates face barriers that UK graduates don't”

Do international medical graduates underperform elsewhere?

This is not, however, a problem unique to the UK. A Canadian study published in 2011 and data in the Medical Council of Canada's *Annual Report 2012* showed that international medical graduates in Canada (including Canadian nationals who had trained outside Canada) underperform in the College of Family Physicians of Canada's and the Royal College of Physicians and Surgeons of Canada's postgraduate examinations. This underperformance is of a similar magnitude to PLAB candidates' underperformance in the MRCP (UK) and the MRCGP compared with UK qualified doctors.

The situation in Canada is despite entry equivalence through the Medical Council of Canada Qualifying Examination – which both Canadian graduates and doctors who trained outside Canada must pass to be able to practise in Canada.

The 2011 study identified the need for better orientation and support for international medical graduates in Canada as well as better training for those who train international medical graduates. These messages chime with findings on the need for better support for doctors new to UK practice in the GMC's report, *The state of medical education and practice in the UK 2011*.*

* General Medical Council (2011) *The state of medical education and practice in the UK 2011* London, GMC.

What is the purpose of the PLAB test?

The purpose of the PLAB test is not to identify whether candidates have the clinical ability and professional attributes required either for postgraduate medical education and training or for progression in the workplace. It does not aim to identify whether candidates have the potential to achieve the equivalent outcomes as UK graduates at this higher level or through medical career pathways. The purpose of the PLAB test is to identify whether candidates have at least the same level of knowledge and skills as the notionally minimally competent UK graduate who has completed F1 training.

In other words, the GMC uses the PLAB test to identify whether candidates demonstrate equivalent outcomes as UK graduates at the point of entry to the register. The standard is set at this level because the GMC considers that it would be unfair to require candidates to demonstrate a higher level of knowledge and skills than UK graduates for the purposes of full registration.

“ The GMC uses the PLAB test to identify whether candidates demonstrate equivalent outcomes as UK graduates”

Why is there a difference in outcomes for PLAB candidates and UK graduates?

The researchers' findings on differential outcomes for PLAB candidates in specialty examinations and postgraduate training do not in themselves amount to evidence either that successful PLAB candidates are not safe to practise in the UK at the level of entry to F2 training or that the pass marks for both Part 1 and Part 2 are set too low. As the Canadian evidence suggests, entry equivalence is not the solution to solving differential outcomes for international medical graduates in postgraduate examinations and training.

We were concerned by the evidence of disparities in outcomes for PLAB candidates compared with UK graduates, but the reasons are far from clear. The researchers suggest significant increases to the pass mark for both parts of the PLAB test with a view to ensuring equivalence in outcomes in the MRCP (UK), the MRCP and the ARCP process.

Recommendation:

The GMC and those responsible for education and training should investigate further why PLAB candidates underperform in the MRCP (UK), the MRCP and the ARCP process compared with UK graduates and the GMC should consider any changes that might be required to the purpose and standard of the PLAB test.

The research findings can't be ignored but the issues are both complex and not unique to the UK. We were unable to recommend a definitive, appropriate and proportionate course of action as the reasons for PLAB candidates' underperformance in the MRCP (UK), the MRCPGP and the ARCP relative to their UK peers appear to be complex.

The effect of increasing the IELTS score

In June 2014, the GMC increased its IELTS overall score requirement from 7.0 to 7.5.

Recommendation:

The GMC should consider further increasing its IELTS score requirements in light of the Durham report's findings on the predictive validity of IELTS for ARCP outcomes.

Maintaining confidence in the PLAB test

The findings of our review broadly fall into the following recommendations.

- Extending the scope of the PLAB test.
- Improving the reliability of the Part 2 examination.
- Limiting attempts at the PLAB test.
- Reducing the currency of passes in the PLAB test.

Taken together, these recommendations should help to make sure that confidence in the PLAB test continues, pending a decision on whether to introduce a UK national licensing examination.

Recommendation:

The GMC should publish a progress report on taking forward work arising from our recommendations in 12–18 months.

Terms of reference and membership

Purpose

The purpose of the review will be to ensure that the PLAB test continues to command the confidence of all key interest groups by being an objective, fair, non-discriminatory and fit-for-purpose method of assessing the knowledge and skills of international medical graduates applying for registration with the GMC.

Key tasks

The key tasks of the review are to consider and make recommendations to Council in relation to the following broad themes:

Theme 1: Ensuring standards

To review whether the knowledge and skills demonstrated by a pass in both parts of the PLAB test continue to be equivalent to those of doctors who have successfully completed F1 training.

To examine whether there should be a limit on the number of attempts at passing the test in order to ensure public confidence in the standards of practice in the UK.

To review whether the periods of validity of passes in both parts of the PLAB test are appropriate.

To consider the feasibility of using the PLAB test for the purpose of a national examination towards the end of F1, in the event that Council were in the future to decide that a national assessment was required at that stage of training.

Theme 2: Content

To review whether the blueprint (which sets out the scope and content of the PLAB test in terms of the topics, skills and procedures) is consistent with the outcomes for provisionally registered doctors described in *The Trainee Doctor* (which replaces *The New Doctor*).

To consider the extent to which the PLAB test could (and should) further assess knowledge and application of the values and principles of *Good medical practice*.

Theme 3: Confidence

To review the reliability and validity of the PLAB test.

This will include examining whether:

- the scoring systems and standard setting reflect best practice
- the Part 1 written examination and Part 2 objective structured clinical examination stations are fit for purpose and reflect current examination and assessment best practice
- the test reliably and accurately differentiates between candidates who meet the required standards of both parts of the test and those who do not.

To review whether the PLAB test and its administration remain compliant with the requirements of the *Equality Act 2010*. This will include the processes for appointing panel members and examiners and for dealing with complaints from candidates.

Theme 4: Outcomes

To examine whether international medical graduates granted full registration following a successful pass in the PLAB test are more or less likely than other cohorts of doctors to experience difficulties in medical practice in the UK.

This will include:

- Examining existing research on the workplace experiences of graduates making the transition to the workplace.
- Examining any evidence of disparity between the success rates of UK medical graduates and those of international medical graduates in postgraduate examinations and assessments.
- Examining the implications of any further research required (as above) as part of the review.

Membership

Professor Kathy Boursicot – Assistant Dean for Medical Education Research and Assessment, Lee Kong Chian School of Medicine, Singapore

Professor Kenneth Cochran* – Chair of the PLA Board

Professor Ian Cumming OBE (Chair) – Chief Executive, Health Education England

Professor Jane Dacre† – GMC Council member (medical)

Dr Sue Davison‡ – GMC Council member (lay)

Professor Derek Gallen – Postgraduate Dean for the Wales Deanery, National Director of the UK Foundation Programme Office, Chair of COPMeD

Dr Alison Graham – Executive Medical Director, NHS Ayrshire and Arran

Ms Roswyn Hakesley-Brown CBE – former Chair of Board of Trustees, Patients Association

Dr Abrar Hussain§ – consultant psychiatrist

Professor Chris McManus – Professor of Psychology and Medical Education, University College London

* Professor Cochran joined the working group when GMC Council members' membership ended on 31 December 2012.

† Professor Dacre's membership of the working group ended when her term of office as GMC Council member ceased on 31 December 2012.

‡ Dr Davison's membership of the working group ended when her term of office as GMC Council member ceased on 31 December 2012.

§ Dr Hussain passed the PLAB test in 2004 and has been registered with the GMC since 2005.

List of recommendations

- The GMC should extend the scope of the PLAB test to include an assessment of the wider ethical values and principles in Good medical practice that the current format can't test.
- The GMC should explore how best to extend the scope of the PLAB test to include an assessment of the professional values and principles in Good medical practice that aren't currently tested. For example, this could be done by including a situational judgement test or another mechanism.
- The GMC should give greater prominence to the PLAB test's blueprint and its overarching statement to improve general understanding of the current scope of the PLAB test.
- The GMC should change the name of the PLAB test to reflect more accurately its purpose. For example, it could be called the GMC's Knowledge and Clinical Skills Tests.
- The GMC should promote the PLAB test's purpose with employers to increase their understanding that the test is set at the level of entry to F2 training.
- The GMC should impose a limit of four attempts at both Part 1 and Part 2 of the PLAB test. Further attempts should only be allowed if circumstances beyond candidates' control have affected performance or on the basis of demonstrable remediation over a period acceptable to the GMC.
- The GMC should provide more meaningful feedback on performance to PLAB test candidates.
- The GMC should require candidates to pass Part 2 of the PLAB test within two years of passing Part 1. It should also require candidates to apply for GMC registration and a licence to practise within two years of passing Part 2.
- The GMC and those responsible for medical education and training in the UK should agree the most appropriate way of delivering a national licensing examination when the future shape of education and training for medical students and newly qualified doctors and the purpose and objectives of the examination are settled.

-
- The GMC should retain the Angoff and borderline group scoring methods for setting the standard of the Part 1 examination and the Part 2 assessment.
 - The GMC should consider using item response theory and statistical equating to support the current standard setting and scoring methodologies for the PLAB test.
 - The GMC should regularly review the PLAB test's standard setting and scoring methodologies to ensure that the test remains up to date and in line with evidence of other methods in a medical context.
 - The GMC should explore the feasibility of introducing computer-based testing for the Part 1 examination and electronic marking for both parts of the PLAB test. This would enable the GMC to provide quicker feedback for candidates, use a wider and even more realistic range of assessment techniques, and gather additional intelligence for analysis.
 - In the interests of patient safety, the GMC should continue to apply one SEM to the pass mark for the Part 2 examination.
 - The GMC should explore how to reduce the possibility of candidates compensating between practical and examination skills stations on the one hand and communication and history taking stations on the other hand. For example, the GMC should consider whether to require a passing standard of performance in the practical and examination skills stations and also in the communication and history taking stations.
 - The GMC should seek to increase the reliability of the Part 2 examination. Options include increasing the number and/or length of the OSCE stations, introducing a feedback session after each OSCE for examiners to discuss candidates' performance, and using two examiners to assess performance. However, the GMC will need to take into account the results of the generalisability study as well as feasibility when determining the most appropriate and proportionate way of increasing the Part 2 reliability.

-
- The GMC should review its procedure for dealing with candidates' requests for reasonable adjustments and its candidate guidance to make sure both reflect relevant learning from the 2012 health and disability review.
 - The GMC should regularly review appeal outcomes to identify and disseminate any trends, lessons learnt and development work required.
 - The GMC should seek to increase the number of female examiners through the continued use of targeted recruitment campaigns.
 - The GMC provides sufficient information to help candidates prepare for the PLAB test and directs them appropriately to other organisations. However, the GMC should regularly review candidates' feedback to make sure that the information and guidance it provides adequately meet their needs.
 - The GMC and those responsible for education and training should investigate further why PLAB candidates underperform in the MRCP (UK), the MRCP and the ARCP process compared with UK graduates and the GMC should consider any changes that might be required to the purpose and standard of the PLAB test.
 - The GMC should consider further increasing its IELTS score requirements in light of the Durham report's findings on the predictive validity of IELTS for ARCP outcomes.
 - The GMC should publish a progress report on taking forward work arising from our recommendations in 12–18 months.

Glossary of terms

Angoff

A method of standard setting based on group judgements about the performance of hypothetical borderline candidates (those who just pass). The GMC uses Angoff to set the pass mark for each Part 1 examination.

ARCP

Annual Review of Competence Progression is the process that scrutinises the suitability of doctors in training to progress to the next stage of, or to complete, a training programme. It is usually held annually, but some specialties have more frequent reviews in the early years of training. Foundation Programmes have incorporated an ARCP since 2013.

Borderline group scoring

The GMC uses the borderline group scoring method to set the pass mark for each station in the Part 2 examination. Examiners make an overall judgement as to whether candidates' performance in each station rates as pass, borderline or fail. The overall judgement is used to determine the pass mark for future candidates. The GMC takes the mean scores of previous candidates judged borderline in each station to work out the pass mark.

Decision study

A decision study uses the results of a generalisability study to model the effects of changing aspects of the OSCE. For example, it will help the GMC to answer questions such as:

'What would be the reliability of the OSCE if it had n stations instead of 14?'

'Would two examiners and n stations attain a particular reliability coefficient?'

'How is the SEM affected by decreasing the number of stations?'

'What would be the effect of increasing the length of the OSCE stations?'

Generalisability study

An extension of classical reliability theory and methodology that indicates the magnitude of errors from various specified sources, such as number of items in the assessment, the number of assessors etc. The analysis is used to indicate the reliability of the test and to evaluate the generalisability beyond the specific sample of items, persons and observational conditions that were studied. A decision study uses the results of a generalisability study to model the effects of changing aspects of the OSCE.

IELTS

The International English Language Testing System is a test that measures the language proficiency of people who want to study or work in environments where English is used. It tests the four language skills – listening, reading, writing and speaking. The GMC accepts required scores in the academic version as evidence of overseas doctors' English language proficiency when they apply for registration.

MRCGP

Postgraduate medical qualification in general practice. Success in the MRCGP examination confirms GP trainees meet RCGP membership requirements and completed the assessments required for specialty training for general practice.

MRCP (UK)

Postgraduate medical qualification in general medicine. Success in the MRCP (UK) confirms medical trainees meet the three medical Royal Colleges' membership requirements and completed the assessments required for specialty training for medicine.

OSCE

Objective structured clinical examination – a multistation clinical examination (typically having 15 to 25 stations). Candidates spend a designated time (usually five to ten minutes) at each station demonstrating a clinical skill or competency at each. Stations frequently feature real or (more often) simulated patients. Artefacts such as radiographs, lab reports and photographs are also commonly used.

Part 2 of the PLAB test is an OSCE which has 14 five-minute stations and a rest station. It may also include a pilot station, for example to test a new scenario, but pilot stations don't count towards a candidate's result. All candidates take the same stations during each Part 2 examination and their performance is marked by an examiner in each station.

Part 1 Panel

The Part 1 Panel is responsible for overseeing and developing the Part 1 examination, and maintaining the question bank. The Part 1 Panel has 14 members, who are all medically qualified associates. They are appointed for their knowledge and expertise in question writing, piloting and standard setting. The panel reports to the PLA Board.

Part 2 Panel

The Part 2 Panel is responsible for overseeing and developing the Part 2 examination and maintaining the bank of OSCE scenarios. The Part 2 Panel has 13 members, who are all medically-qualified associates. They are appointed for their knowledge and expertise in running OSCEs and in question writing, piloting and standard setting. The panel also has a communications consultant and reports to the PLA Board.

PLA Board

The PLA Board has overall responsibility for overseeing and developing the PLAB test. The Board has 11 members, who are associates. Nine are medically qualified and have a range of experience in assessment practices, as well as undergraduate and Foundation Programme education and training. They include an international medical graduate who has passed the PLAB test. Two are not medically qualified.

PLAB blueprint

The blueprint is effectively the PLAB test's curriculum. It sets out the scope and content of the test in terms of the topics, skills and procedures that a doctor who passes the test would need to know and be able to do, and the professional qualities expected of a doctor working in the UK.

Reliability

Reliability expresses a trust in accuracy or the provision of the correct results. In the case of assessments, it is an expression of internal consistency and reproducibility. This quality is usually calculated statistically and reported as coefficient alpha, which is a measure of a test's internal consistency. Generalisability theory is becoming the preferred alternative because, although it is considerably more complicated to calculate, it gives much richer information.

A perfectly reproducible test would have a coefficient of 1.0. That is, all candidates would achieve the same score on retesting. In reality, tests are affected by many variables and combinations of factors, so a figure of 1.0 is unrealistic. A reliability coefficient of greater than 0.8 is frequently quoted as an appropriate figure for high-stakes assessments* (although it is acknowledged that reliability measures are difficult or even impossible for small cohorts).

SEM

The standard error of measurement is the standard deviation of errors of measurement that are associated with test scores from a particular group of examinees. It is important in identifying borderline candidates.

Single best answer

An examination format in which each question consists of a stem and a set of possible answers from which the candidate must choose the best answer. The format allows facts to be placed in a clinical context so that the application of knowledge and problem solving can be tested.

Situational judgement tests

Situational judgement tests present realistic hypothetical scenarios and ask candidates to identify an appropriate response, generally in multiple choice format.

Standard deviation

Standard deviation is a widely used measure of variability, showing how much variation or dispersion there is from the average (mean) or expected value. A low standard deviation indicates that the data points tend to be very close to the mean, whereas a high standard deviation indicates that the data points are spread out over a large range of values. The standard deviation is one component of the equation to calculate the SEM.

Validity

In the case of assessment, validity refers to the degree to which a measurement instrument truly measures what it is supposed to measure. It is concerned with whether the right things are being assessed, in the right way, and with a positive influence on learning.

* General Medical Council (2010) *Reliability issues in the assessment of small cohorts* London, GMC.

Email: gmc@gmc-uk.org
Website: www.gmc-uk.org
Telephone: **0161 923 6602**

General Medical Council, 3 Hardman Street, Manchester M3 3AW

Textphone: **please dial the prefix 18001** then
0161 923 6602 to use the Text Relay service

Join our conversation

 [@gmcuk](https://twitter.com/gmcuk)  facebook.com/gmcuk
 linkd.in/gmcuk  youtube.com/gmcuktv

This information can be made available in alternative formats or languages. To request an alternative format, please call us on **0161 923 6602** or email us at publications@gmc-uk.org.

Published September 2014

© 2014 General Medical Council

The text of this document may be reproduced free of charge in any format or medium providing it is reproduced accurately and not in a misleading context. The material must be acknowledged as GMC copyright and the document title specified.

The GMC is a charity registered in England and Wales (1089278) and Scotland (SC037750)

Code: GMC/PLABR/0914

General
Medical
Council