

UK medical selection: lottery or meritocracy?

Authors: Benjamin HL Harris,^A Jason L Walsh^B and Simon Lammy^C

ABSTRACT

From senior school through to consultancy, a plethora of assessments shape medical careers. Multiple methods of assessment are used to discriminate between applicants. Medical selection in the UK appears to be moving increasingly towards non-knowledge-based testing at all career stages. We review the evidence for non-knowledge-based tests and discuss their perceived benefits. We raise the question: is the current use of non-knowledge-based tests within the UK at risk of undermining more robust measures of medical school and postgraduate performance?

KEYWORDS: Medical selection, UKCAT, BMAT, GAMSAT, situational judgement tests

Introduction

From senior school through to consultancy, a plethora of assessments shape medical careers. Selection for medical school, foundation posts and beyond is a difficult and controversial process, in which a large number of seemingly similar applicants compete for a small number of positions.

Therefore, multiple methods of assessment are used to discriminate between applicants. These range from the traditional methods of assessing academic achievement and interview performance, to the relatively recent inclusion of ‘non-knowledge-based’ tests, such as the UK Clinical Aptitude Test (UKCAT) and situational judgement tests (SJT). In this context, we define non-knowledge-based tests as assessments which have been devised to test desirable professional attributes, such as empathy, ethical awareness and logical thinking, not the clinical knowledge/skills of trainee doctors. In this article, we discuss these approaches and ask whether they are appropriate.

In light of the recent proposed changes to medical career structure in the UK,¹ we believe it is timely that medical selection is re-examined. It is our view that the current choice

of selection methods in the UK seems to have adopted a ‘fashion-based’ rather than an evidence-based approach to selecting the best candidates.

Selection for UK medical schools

Selection of candidates for medical school used to be focused purely on school performance, particularly in science, and performance at interview. There were several objections to this approach. First, potential doctors were considered to need additional qualities than those associated with a good scientist and scholar; particularly social attributes, such as empathy and communication skills.² Second, interviews were thought to result in selection of candidates ‘in the image’ of the interviewers, which may have led to a bias towards a particular gender or social class.^{3,4} In an attempt to address this problem and to discriminate between large numbers of academically similar school leavers, several tests with a significant non-knowledge-based component were introduced (Table 1).

Currently, medical schools in the UK use three broad areas to select applicants. First, academic achievement and commitment to medicine are assessed on the Universities and Colleges Application System (UCAS) form. Following this, selection testing is conducted through one of three tests; the UKCAT, Biomedical Medical Admissions Test (BMAT) and Graduate Australian Medical School Admissions Test (GAMSAT) (Table 1). Finally, potential medical students are evaluated in conjunction with their personal statements via traditional or multiple mini-interviews (MMIs).

Great variability exists between medical schools, as some use all three broad areas to evaluate an applicant’s suitability to read medicine, whereas others use only one (eg academic performance).⁵ The majority of medical schools use at least one test with a significant non-knowledge-based component in selection, under the belief that these test intrinsic abilities that are desirable in future practice. Differences in these tests exist in terms of their scientific knowledge content, which is absent in the UKCAT, contributes one-third to the BMAT and one-half to the GAMSAT.

Do non-knowledge-based tests predict good students?

Correlation of non-knowledge-based tests with medical school performance has been questioned on several occasions, but these tests still persist as part of medical school selection.^{6–12} Several recent studies have shown that the UKCAT does not correlate, or correlates only weakly, with performance in medical or dental school.^{7,8,13,14} In the largest study to date

Authors: ^Aacademic foundation doctor, John Radcliffe Hospital, Oxford, UK, and non-stipendiary lecturer, St Anne’s College, University of Oxford, Oxford, UK; ^Bacademic foundation doctor, University Hospital of Wales, Cardiff, UK; ^CNIHR ACF vascular surgery, Nuffield Department of Surgical Sciences, University of Oxford, John Radcliffe Hospital, Oxford, UK, and ST1 neurosurgery, Department of Neurosurgery, Institute of Neurological Sciences, Southern General Hospital, Glasgow UK

Table 1. Non-knowledge-based tests used in undergraduate medical selection in the UK.

Tests with a significant non-knowledge-based component	Description	Duration	Content
UK Clinical Aptitude Test (UKCAT)	'The UKCAT does not contain any curriculum or science content. It focuses on exploring the cognitive powers of candidates and other attributes considered to be valuable for health care professionals.' www.ukcat.ac.uk/	2 hours	Verbal, quantitative and abstract reasoning, decision analysis and situational judgement.
Biomedical Medical Admissions Test (BMAT)	'BMAT is a subject-specific admissions test for applicants to medicine, veterinary medicine and similar courses at universities.' www.admissionstesting-service.org/	2 hours	Aptitude and skills, scientific knowledge and applications, and a writing task (in a scientific theme).
Graduate Australian Medical School Admissions Test (GAMSAT)	'GAMSAT evaluates the nature and extent of abilities and skills gained through prior experience and learning, including the mastery and use of concepts in basic science as well as the acquisition of more general skills in problem solving, critical thinking and writing.' http://gamsat.acer.edu.au/gamsat-uk/	5 hours 30 minutes	Reasoning in humanities and social sciences, written communication and reasoning in biological and physical sciences.

McManus *et al* showed that a weak correlation does exist between UKCAT performance and overall performance in early medical school exams ($r=0.148$; $n=4811$), but this is weaker than the correlation with a variety of measures of educational achievement in secondary education.¹⁵ Therefore, any potential impact of the UKCAT is diminished by measures of prior course performance, such as A-level results. This pattern is seen again and even more evident when examining six longitudinal cohort studies across a range of universities, where A levels are shown as reasonable predictors of performance throughout medical school and beyond (Table 2).¹⁶

One potential advantage of the UKCAT is that it increases the number of individuals from under-represented socioeconomic groups who are offered places at medical school.^{17,18} However, it is unclear whether this represents anything other than increased randomness in its selection from the candidate pool. In other words, it is not clear whether the UKCAT selects the most suitable candidates from under-represented socioeconomic groups.

The BMAT was initially shown to correlate better with medical school performance.¹⁹ However, recent analysis has shown its predictive nature is overwhelmingly attributable to the scientific knowledge component.¹¹ Similarly, the only section of the GAMSAT that has been shown to predict medical school performance is the scientific knowledge section.⁹

Table 2. Construct validity coefficients of A levels versus aptitude tests, adapted from McManus *et al*.¹⁶

Examinations	First year results	Overall undergraduate performance	Overall MRCP results
A levels	0.809	0.723	0.506
Aptitude tests	0.245	0.181	0.226

Despite the lack of evidence for their validity as a reliable selection tool, non-knowledge-based tests are still used in many medical schools as a major component of the selection process.⁵ For example, some medical schools weight the UKCAT as 50% of the overall score to determine a candidate's eligibility for interview. If the UKCAT, BMAT or the GAMSAT are used as the first selection filter, candidates with good A levels are rejected. An alternative consequence is that, since most candidates have very similar A levels, the only differentiating factor is their score on non-knowledge-based test. Therefore non-knowledge-based tests effectively take on the role as the primary selector, despite their poor validity. We question whether it is appropriate to select candidates for medical school with tests which have little predictive value. Surely, we should aim to select applicants using the most robust predictors of medical school performance available.

Another approach to test non-knowledge-based attributes desired in medical students is the use of MMIs.³ MMIs have shown predictive validity in the context of postgraduate medical school admissions in the USA.²⁰⁻²² Recently, the University of Dundee School of Medicine carried out a large prospective analysis of their admission assessments and found that MMIs were a significant predictor of early success in medical school. By contrast, UKCAT scores and scores derived from both references and personal statements did not correlate with early medical school performance.²³ Therefore, there is potentially a place for MMIs in the context of UK medical school admission. However, for a more robust analysis, these findings should be replicated in other UK medical schools and MMI scores should also be correlated with later degree and postgraduate performance in the UK. A potential drawback of MMIs is that they could be logistically complex to organise on a national scale.

A-level performance has been shown to correlate strongly with medical school performance and postgraduate performance.^{15,24,25} Recently, however, its discriminatory ability has decreased with a large proportion of applicants attaining

at least three A grades.²⁶ With addition of A* grades, the hope is to restore the discriminatory power of A levels. However, it is not certain that this is a sufficient change. Past experience shows that A-level grade inflation is inevitable. In addition, the existence of different examination boards, with potentially different standards, makes A-level grades difficult to interpret unequivocally.

The literature indicates that it might be more prudent to use A-level grades as a primary selector of candidates, as they are a validated predictor of medical school performance.²⁵ However, since so many candidates achieve top grades in their A levels, scores on tests with significant non-knowledge-based components assume a disproportionate importance in selection, despite their poor validity. We therefore need an evidence-based alternative.

In the USA, medical students are selected on their undergraduate grade point average and the national Medical College Admissions Test (MCAT). The MCAT was introduced in 1946 and has been reviewed and reshaped on several occasions. The current version consists of three sections; biological sciences, physical sciences and verbal reasoning. The MCAT has been shown to be a strong predictor of academic performance at medical school and in all three steps of the United States Medical Licensing Examinations.^{27,28} The most predictive parts of the MCAT are the biological and physical sciences sections, which test scientific knowledge.^{27,28} Taken in conjunction with the evidence that scientific components from other tests are the best predictors of medical school performance,^{11,12} it would seem appropriate that medical student selection should be focused around testing scientific knowledge and reasoning.

It has been repeatedly demonstrated that tests of a student's ability to learn and use information tend to predict medical school performance, whereas exams testing abstract qualities have a limited value.^{7,9,15} Therefore we believe testing abstract qualities, such as quantitative reasoning and situational judgement, should not be included, or at least weighted very lightly until they are validated. We believe the current focus on non-knowledge-based tests risks undervaluing proven markers of future performance. However, A-level grades are not sufficiently discriminatory to select a subset of medical students from a host of high-performing school leavers. Furthermore, in designing a new test, an evidence-based approach should be adopted. If the test is ineffective, the evidence should not be ignored and the test reformed. On this basis, we propose a national scientific knowledge-based test to be taken in conjunction with A levels, for initial selection into medical school – the UK Pre-Medical Exam (UKPME).

UKPME

Candidates are given a test containing relevant bridging information from sixth form to medical school. This test will include material on the basic sciences, including anatomy, physiology and biochemistry, and should include some scientific problem-based analysis. We propose forming a committee of academics, teachers and senior clinicians to put together a specification for this test, which should incorporate such material together with A-level content. The aim is that students will focus their learning on material relevant to medical school and so applicants will be better equipped when

starting their course, rather than spending an inordinate amount of time and effort trying to perfect their technique for abstract selection tests. The Medical Schools Council could lead the implementation of the UKPME.

We propose a multiple-choice question (MCQ) format for the UKPME. The reasons for employing MCQs include:^{29,30}

- > MCQs test a large amount of material in a short amount of time
- > MCQs are easy to evaluate for validity and reliability
- > papers can be electronically marked and therefore do not require any subjective input.

The mark on the UKPME will be shared with the relevant medical schools to be used in the admissions process.

Foundation Programme selection

In the final year of medical school, candidates apply for junior doctor positions in the UK Foundation Programme. This application is based on academic achievement (the educational performance measure (EPM)) and performance on a situational judgement test (SJT). The EPM is a score derived from the candidate's ranking in their respective medical school (decile), additional degrees and publications. The maximum score on this element is 50 points, but the minimum score is 34 points, assuming the candidate completes a medical degree. Therefore this discriminates effectively on a 16-point scale.

The SJT is a test lasting 2 hours and 20 minutes which consists of a series of questions containing hypothetical scenarios that junior doctors may encounter. Candidates are asked to rank the appropriateness of actions or pick the most appropriate three actions from a list. The SJT contributes a maximum of 50-points towards the total job application score. The minimum score for the SJT is 0 points and so the SJT is measured on a 50 point scale (although candidates rarely score below 30).³¹

There are several problems with this approach, which, in part, are reminiscent of the problems of medical school selection; the predominant issue is that non-knowledge-based tests are becoming a major component of selection with minimal evidence.

Problems with Foundation Programme selection

The EPM

The EPM depends upon a candidate's ranking in their respective medical school (decile). However, the points distribution system, based on academic decile, may be flawed. The proportion of high-achieving students at different medical schools may vary, and the way students are marked at medical school is not standardised between institutions. So, in effect, a student in the fifth decile in (highly competitive) university A may be equivalent to one in the second decile in (less competitive) university B. It has been demonstrated that students from different medical schools perform with varying abilities in postgraduate examinations.^{32,33}

The SJT

SJTs do not have any proven efficacy in selecting appropriate candidates for junior doctor positions. In medicine, the validity of the SJT has been extensively evaluated only in the specific context of UK General Practice selection,³⁴ and low but significant correlations between SJT scores and measures of

General Practice performance have been found. It is difficult to assess the validity of these low correlations, as the study did not provide data on the distribution of individual scores. Furthermore, these correlations are with context-specific measures and therefore cannot be readily generalised to all of UK medical selection.

There is no guarantee that SJTs measure ‘personality and implicit trait policies’ as they intend.³⁵ Indeed, Koswara *et al* showed that there are correlations between SJT scores, intelligence quotient (IQ) tests and clinical knowledge test scores, suggesting that all three tests assess overlapping constructs.³⁵ In addition, there are no data suggesting candidates who score highly on SJTs actually behave more appropriately in medicine and there is no evidence that unprofessional individuals without the desired abstract characteristics score poorly on SJTs. Written SJTs are low fidelity simulations in which individuals can answer what is perceived as the correct option. This does not necessarily reflect what an individual’s actions would be in a real situation. In fact, Nyugen *et al* showed that ‘SJTs can be faked...[whereas] Knowledge response [items] are more resistant to faking’.³⁶

The discriminatory ability of the SJT is consistently poor in Foundation Programme selection. The 2012–13 and 2013–14 SJTs, and the preceding SJT pilots showed similar results: candidates’ scores tended to cluster close to the mean. In the distribution of raw SJT scores in 2014, the mean mark was 836.4 and the standard deviation (sd) was 27.9.³¹ This has led to small differences in the SJT raw score being extrapolated into a large points difference in this and previous years.³⁷ Furthermore, the scores were so closely clustered that missing out three questions of the 60 ‘live’ questions (eg due to time constraints) could move a candidate from the average into the bottom 3% of applicants.^{31,38} In addition, students who guess randomly on one question would score 1.85 scaled marks (0.44 sd) more than ones who do not answer that question at all, with presumably, no extra competence (see Box 1). To give an idea of importance of 1.85 marks on the job application scale, a 2:1 degree is awarded 3 points on the EPM (Table 3).

Early face validity measures of the SJT in the context of Foundation Programme are poor. According to the Working Party Group: ‘only 38.6% agreed or strongly agreed that the content of the SJT appeared to be fair for selection to the foundation programme...[while only] 25.4% of applicants agreed or strongly agreed that the results of the SJT should help selectors to differentiate between weaker and stronger applicants’.³⁸ Similarly poor results are seen in the 2014 analysis.³¹

Currently, the foundation SJT is weighted as 50% of the total job application score in Foundation Programme selection.³⁹ However, we have shown it is a volatile measure of competency and has various other drawbacks. It is a worrying trend that SJTs are now being incorporated into selection for medical school and postgraduate posts.^{15,35,40}

By contrast, it has been shown that medical school performance and scores on a national knowledge-based medical examination correlate well with postgraduate examination performance.^{41–44} If an SJT is to be used as part of any assessment process, it should be a low-stakes component of the overall assessment until it is validated in context. The current selection process risks diverting the attentions of final year medical students from perfecting clinical examination and learning applicable medical knowledge to ‘studying’ for the SJT.

Box 1. Why guessing pays dividends.

Format:

SJT paper = 70 questions in 2 hours 20 minutes

60 questions are ‘live’ and 10 ‘pilot’ questions are not included in the participant’s final score

Approximately two-thirds of questions are ranking questions and one-third are ‘select three from eight’ MCQs.

Job application score scaling method:

SJT converted score = (SJT raw score^A x 0.154) – 90.1

Formula from the UK Foundation Programme Office

Ranking question:

- > Maximum raw score per question = 20/20 marks
 - (20 x 0.154) = 3.08 job application marks
- > Most likely score if guessed = 12/20
 - (12 x 0.154) = 1.848 job application marks
- > No attempt = 0 job application marks

‘Select three from eight’ MCQs:

- > Maximum raw score per question = 12/12 marks
 - (12 x 0.154) = 1.848 job application marks
- > Average raw score if guessed = 4.5/12
 - (4.5 x 0.154) = 0.693 job application marks
- > Minimum score per attempt/no attempt = 0/12

^AStrictly test equated raw score across sittings.

Available online at www.foundationprogramme.nhs.uk/pages/home/how-to-apply/SJT-FAQs#answer379K

MCQ = multiple choice questions; SJT = situational judgement test.

Table 3. Points attributed to postgraduate degrees in Foundation Programme selection.³⁹

Points awarded (maximum = 5)	Degree category
0	Primary Medical Qualification only 3rd class BMedSci awarded at University of Nottingham
1	3rd class honours degree Unclassified honours degree 2.2 class BMedSci awarded at University of Nottingham
2	2.2 class honours degree 2.1 class BMedSci awarded at University of Nottingham
3	2.1 honours degree 1st class BMedSci awarded at University of Nottingham
4	Postgraduate Masters degree 1st class honours degree Bachelor of Dental Surgery (BDS) Bachelor Veterinary Medicine (B Vet Med)
5	Doctoral degree (PhD, DPhil, etc)

A solution: the UKNLE

Following these considerations, we propose that there should be a national examination across the UK for Foundation Programme jobs, where all candidates can be assessed against a single standard: the UK National Medical Licensing Examination (UKNLE). This centralised final examination should be knowledge-based and examinations could be generated from a central database of questions. The examination could be staged and taken when students have covered the desired central curriculum material. The timing of when students have covered this material may vary between medical schools due to differences in individual curricula, but a medical school could recommend to its students the appropriate time to undertake the national examination (eg after third year and mid-way through final year).

This approach does not preclude individual medical schools from setting their own examinations to particular local standards (eg for award of honours) if desired.

The UK Foundation Programme Office in conjunction with the General Medical Council and Medical Schools Council are well placed to implement the UKNLE.

Postgraduate training

In the past decade, the landscape of postgraduate medical and surgical training has drastically changed. It has evolved from an in-house, non-standardised and highly geographically-dependent selection process towards a national, standardised and highly reproducible scheme. Nonetheless, assessment has continued to emphasise testing the core competencies directly, either through examination or by judging published output. We are concerned that future developments in assessment may move away from this 'academic' approach towards non-knowledge-based testing, as has been the case for selection procedures in earlier stages of progression.

Currently, junior doctors intent on negotiating selection into core specialties, eg medicine and surgery, and even some of the run-through specialties, eg neurosurgery, must demonstrate increasing commitment to acquire the specialty, post and hospital of their choice.⁴⁵ Selectors assess this commitment across a number of domains: clinical, academic, managerial and professional achievement, and aptitude. For most specialties, demonstrations of clinical commitment stretch far beyond acquiring the basic competencies in the Foundation Programme curriculum, and include completing several postgraduate clinical courses. Some courses are broad, for example, Advanced Life Support and Advanced Trauma Life Support. However, most courses are specialty-specific, for instance, 'Care of the Critically Ill Neurosurgical Patient' for neurological surgery, 'Focused Echocardiography in Emergency Life Support' for core anaesthesia, and 'Ill Medical Patients' Acute Care and Treatment' for core medical and general practice. A significant milestone in progression is the passing of postgraduate examinations (eg membership of the Royal College of Surgeons). These results provide a robust and well-respected marker of academic achievement and potential. Currently, these courses and examinations are taken together with an assessment of research performance (as assessed by published papers) and performance in clinical audit, to assess suitability for advanced training programmes.

One might argue that this approach puts excessive pressure on junior doctors, attempting to negotiate an ever-expanding extracurricular workload. The European Working Time Directive has attempted to lighten this load by restricting time spent on the wards, but the effect has been to increase expectations of publication, and hence time spent on 'extra-clinical activities' by trainees. Publication rates by junior surgical trainees have indeed increased greatly in the last few years.⁴⁶ The pressure to publish, to complete these advanced courses, as well as fulfilling basic clinical competences may be seen as becoming excessive, and may need to be reconsidered in the future. Nonetheless, the principle that selection should be heavily weighted on academic performance has been maintained, and, if the workload balance can be restored, would be considered by many to be the most reliable and desirable method. However, the predictive validity of research performance in postgraduate selection is still uncertain.

Until recently, the above competencies were assessed using structured forms for both core and specialty training post selection, in addition to an interview. At interview, candidates are assessed across several stations:

- > portfolio station, to permit an oral discussion of commitment to a particular career path
- > several clinical stations, to assess clinical decision making, management and lateral thinking
- > management station, to analyse non-clinical management experience, expertise and leadership potential
- > communication station, to evaluate higher order communications skills with both simulated patients and colleagues
- > skills stations particular to the specialty.

The various specialties have typically selected candidates using a variety of the above stations (varying in number, length and complexity) appropriate to the specialty and experience of the candidate. It is noticeable, however, that an increasing number have piloted the SJT in previous selection rounds, and that the 2014/2015 applications for specialty saw a growing number of specialties opting to use SJTs as part of the selection process.

This raises the concern that the expanding extracurricular workload now expected of junior doctors is at risk of being devalued in the same way as undergraduate medical school performance. It seems that, progressively, the vertical spectrum of medical selection, from entry into medical school, the Foundation Programme and core and specialty training, is being readjusted to replace measures of relevant medical knowledge with non-knowledge-based tests.

Conclusion

Medical selection in the UK appears to be moving increasingly towards non-knowledge-based testing at all career stages. Currently, the selection process places more emphasis on non-knowledge-based testing than is justified by evidence. The current favourite, SJTs, are already a significant factor in the framework of medical selection despite limited evidence for their value. If non-knowledge-based tests continue to proliferate in postgraduate selection, they may persist and become overvalued, as they have become for medical school selection (UKCAT, BMAT and GAMSAT). We feel it is timely to argue against this possibility. In addition, the weighting of

abstract tests of this type in selecting for career progression may well discourage doctors, at all career stages, from acquiring medical knowledge and pursuing research – both skills which are highly valued in medicine.

We understand the pressures to test ‘other qualities’ of doctors and to ensure that clinicians are safe. However, there is little evidence that desired qualities are assessed by non-knowledge-based tests. There is certainly no evidence that these tests produce safer doctors. On the contrary, national knowledge-based tests, such as the MCAT in the USA, show a much stronger correlation with medical licensing examination performance.²⁷

Safe clinicians are doctors who have perfected the art of clinical examination and who are able to interpret and synthesise complex information, while still caring for the patient in front of them. To do this, knowledge and clinical acumen are key. Therefore, let us halt this fashion-based approach to medical selection and focus our attention on the evidence, and what really matters, diagnosis, management and patient care. ■

Acknowledgements

We thank Professor David Wilson for his helpful comments in the preparation of this article.

References

- Greenway D. *Securing the future of excellent patient care: shape of training review*. Manchester: Shape of Training, 2013.
- General Medical Council. *Tomorrow's Doctors (2009)*. Manchester: GMC, 2009.
- Albanese MA, Snow MH, Skochelak SE, Huggett KN, Farrell PM. Assessing personal qualities in medical school admissions. *Acad Med* 2003;78:313–21.
- Edwards JC, Johnson EK, Molitor JB. The interview in the admission process. *Acad Med* 1990;65:167–77.
- Adam J, Dowell J, Greatrix R. Use of UKCAT scores in student selection by U.K. medical schools, 2006–2010. *BMC Med Educ* 2011;11:98.
- Cassidy J. UKCAT among the pigeons. *BMJ* 2008;336:691–2.
- Lynch B, Mackenzie R, Dowell J, Cleland J, Prescott G. Does the UKCAT predict Year 1 performance in medical school? *Med Educ* 2009;43:1203–9.
- Yates J, James D. The UK Clinical Aptitude Test and clinical course performance at Nottingham: a prospective cohort study. *BMC Med Educ* 2013;13:32.
- Groves MA, Gordon J, Ryan G. Entry tests for graduate medical programs: is it time to re-think? *Med J Aust* 2007;186:120–3.
- McManus C, Powis D. Testing medical school selection tests. *Med J Aust* 2007;186:118–9.
- McManus IC, Ferguson E, Wakeford R, Powis D, James D. Predictive validity of the Biomedical Admissions Test: an evaluation and case study. *Med Teach* 2011;33:53–7.
- Wilkinson D, Zhang J, Byrne, GJ *et al*. Medical school selection criteria and the prediction of academic performance. *Med J Aust* 2008;188:349–54.
- Lala R, Wood D, Baker S. Validity of the UKCAT in applicant selection and predicting exam performance in UK dental students. *J Dent Educ* 2013;77:1159–70.
- Sartania N, McClure JD, Sweeting H, Browitt A. Predictive power of UKCAT and other pre-admission measures for performance in a medical school in Glasgow: a cohort study. *BMC Med Educ* 2014;14:116.
- McManus IC, Dewberry C, Nicholson S, Dowell JS. The UKCAT-12 study: educational attainment, aptitude test performance, demographic and socio-economic contextual factors as predictors of first year outcome in a cross-sectional collaborative study of 12 UK medical schools. *BMC Med* 2013;11:244.
- McManus IC, Dewberry C, Nicholson S *et al*. Construct-level predictive validity of educational attainment and intellectual aptitude tests in medical student selection: meta-regression of six UK longitudinal studies. *BMC Med* 2013;11:243.
- Lambe P, Waters C, Bristow D. The UK Clinical Aptitude Test: is it a fair test for selecting medical students? *Med Teach* 2012;34:557–65.
- Tiffin PA, Dowell JS, McLachlan JC. Widening access to UK medical education for under-represented socioeconomic groups: modelling the impact of the UKCAT in the 2009 cohort. *BMJ* 2012;344:1805.
- Emery JL, Bell JF. The predictive validity of the BioMedical Admissions Test for pre-clinical examination performance. *Med Educ* 2009;43:557–64.
- Eva KW, Reiter HI, Rosenfeld J, Norman GR. The ability of the multiple mini-interview to predict pre-clerkship performance in medical school. *Acad Med* 2004;79(10 Suppl):S40–2.
- Eva KW, Reiter HI, Trinh K *et al*. Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ* 2009;43:767–75.
- Reiter HI, Eva KW, Rosenfeld J, Norman GR. Multiple mini-interviews predict clerkship and licensing examination performance. *Med Educ* 2007; 41:378–84.
- Husbands A, Dowell J. Predictive validity of the Dundee multiple mini-interview. *Med Educ* 2013;47:717–25.
- McManus IC, Powis DA, Wakeford R *et al*. Intellectual aptitude tests and A levels for selecting UK school leaver entrants for medical school. *BMJ* 2005;331:555–9.
- McManus IC, Smithers E, Partridge P, Keeling A, Fleming PR. A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. *BMJ* 2003;327:139–42.
- McManus C, Woolf K, Dacre JE. Even one star at A level could be “too little, too late” for medical student selection. *BMC Med Educ* 2008;8:16.
- Callahan CA, Hojat M, Veloski J, Erdmann JB, Gonnella JS. The predictive validity of three versions of the MCAT in relation to performance in medical school, residency, and licensing examinations: a longitudinal study of 36 classes of Jefferson Medical College. *Acad Med* 2010;85:980–7.
- Julian ER. Validity of the Medical College Admission Test for predicting medical school performance. *Acad Med* 2005;80:910–7.
- Case SM, Swanson DB. *Constructing written test questions for the basic and clinical sciences*. Philadelphia: National Board of Medical Examiners, 2000.
- Hayes R. Assessment in medical education: roles for clinical teachers. *Clin Teach* 2008;5:23–7.
- Patterson F, Murray H, Baron H, Aitkenhead A, Flaxman, C. *Analysis of the Situational Judgement Test for selection to the Foundation Programme 2014*. Derby: Work Psychology Group, 2014.
- McManus IC, Elder TA, de Champlain A *et al*. Graduates of different UK medical schools show substantial differences in performance on MRCP (UK) Part 1, Part 2 and PACES examinations. *BMC Med* 2008;6:5.
- Rushd S, Landau AB, Khan JA, Allgar V, Lindow SW. An analysis of the performance of UK medical graduates in the MRCOG Part 1 and Part 2 written examinations. *Postgrad Med J* 2012; 88:249–54.
- Patterson F, Lievens F, Kerrin M, Munro N, Irish B. The predictive validity of selection for entry into postgraduate training in general practice: evidence from three longitudinal studies. *Br J Gen Pract* 2013;63:734–41.
- Koczwara A., Patterson F, Zibarras L *et al*. Evaluating cognitive ability, knowledge tests and situational judgement tests for postgraduate selection. *Med Educ* 2012;46:399–408.
- Nguyen NT. Effects of response instructions on faking a situational judgment test. *IJSA* 2005;13:250–60.
- Walsh JL, Harris BHL. Situational Judgement Tests: more important than educational performance? *BMJ Careers*, 2013. Available online at <http://careers.bmj.com/careers/advice/view-article.html?id=20011842> [Accessed 13 October 2014].

- 38 Patterson F, Ashworth V, Murray H, Empey L, Aitkenhead A. *Analysis of the Situational Judgement Test for selection to the Foundation Programme 2013*. Derby: Work Psychology Group, 2013.
- 39 UK Foundation Programme Office. *FP/AFP 2014 applicants' handbook*. Cardiff: UK Foundation Programme Office, 2013.
- 40 Carr A, Irish B. The new specialty selection test. *BMJ Careers*, 2013. Available online at <http://careers.bmj.com/careers/advice/view-article.html?id=20014944> [Accessed 13 October 2014].
- 41 McManus IC, Woolf K, Dacre J, Paice E, Dewberry C. The Academic Backbone: longitudinal continuities in educational achievement from secondary school and medical school to MRCP(UK) and the specialist register in UK medical students and doctors. *BMC Med* 2013;11:242.
- 42 Dougherty PJ, Walter N, Schilling P, Najibi S, Herkowitz H. Do scores of the USMLE Step 1 and OITE correlate with the ABOS Part I certifying examination?: a multicenter study. *Clin Orthop Relat Res* 2010;468:2797–802.
- 43 Perez JA, Greer S. Correlation of United States Medical Licensing Examination and Internal Medicine In-Training Examination performance. *Adv Health Sci Educ Theory Pract* 2009;14:753–8.
- 44 Fening K, Vander Horst A, Zirwas M. Correlation of USMLE Step 1 scores with performance on dermatology in-training examinations. *J Am Acad Dermatol* 2011;64:102–6.
- 45 Carr A, Marvell J, Collins J. Applying to specialty training: considering the competition. *BMJ Careers*, 2013. Available online at <http://careers.bmj.com/careers/advice/view-article.html?id=20015362> [Accessed 13 October 2014].
- 46 Ansell JM, Beamish AJ, Warren N, Torkington J. How do surgical trainees without a higher degree compare with their postdoctoral peers? *Bulletin of The Royal College of Surgeons of England* 2013;95:1–5.

**Address for correspondence: Dr BHL Harris, St Anne's College, Woodstock Road, Oxford OX2 6HS, UK.
Email: benjamin.harris@medsci.ox.ac.uk**