Research

Fiona Patterson, Filip Lievens, Máire Kerrin, Neil Munro and Bill Irish

The predictive validity of selection for entry into postgraduate training in general practice:
evidence from three longitudinal studies

Abstract

Background
The selection methodology for UK general practice is designed to accommodate several thousand applicants per year and targets six core attributes identified in a multi-method job-analysis study

Aim
To evaluate the predictive validity of selection methods for entry into postgraduate training, comprising a clinical problem-solving test, a situational judgement test, and a selection centre.

Design and setting
A three-part longitudinal predictive validity study of selection into training for UK general practice.

Method
In sample 1, participants were junior doctors applying for training in general practice (*n* = 6824). In sample 2, participants were GP registrars 1 year into training (n = 196). In sample 3, participants were GP registrars sitting the licensing examination after 3 years, at the end of training (*n* = 2292). The outcome measures include: assessor ratings of performance in a selection centre comprising job simulation exercises (sample 1); supervisor ratings of trainee job performance 1 year into training (sample 2); and licensing examination results, including an applied knowledge examination and a 12-station clinical skills objective structured clinical examination (OSCE; sample 3).

Results
Performance ratings at selection predicted subsequent supervisor ratings of job performance 1 year later. Selection results also significantly predicted performance on both the clinical skills OSCE and applied knowledge examination for licensing at the end of training.

Conclusion
In combination, these longitudinal findings provide good evidence of the predictive validity of the selection methods, and are the first reported for entry into postgraduate training. Results show that the best predictor of work performance and training outcomes is a combination of a clinical problem-solving test, a situational judgement test, and a selection centre. Implications for selection methods for all postgraduate specialties are considered.

Keywords
general practice; postgraduate; reliability and validity; student selection.

INTRODUCTION

Selection into postgraduate medical training has been a relatively under-researched topic[1–3] and, as with any selection methodology, various psychometric and legal criteria must be satisfied, including standardisation, reliability, validity, and fairness.[4–6] This study presents evidence from a three-part longitudinal study examining the predictive validity of a selection system used to appoint trainees into postgraduate training in the UK, building on a previous initial validation study,[7] linking selection data with subsequent in-training assessments and training outcomes. Although there is evidence emerging on the predictive validity of selection methods for medicine,[8] and in exploring demographic and educational factors associated with licensure certification,[9] to the authors' knowledge, this is the first study to report on a large scale the long-term predictive validity of postgraduate selection methods. Recently, there has been much debate on medical schools' admission processes, where policies vary internationally.[1] Faced with limited training posts and large numbers of applicants, most recruiters have traditionally relied on academic criteria in admission procedures. Almost universally, high academic achievement is a minimum entry requirement for medical school admissions, which assumes that with good academic ability, the other skills required to be a competent clinician are then trainable. This study presents data to encourage further debate and to develop a future international research agenda for postgraduate selection, drawing implications for design of the selection system in general.

The selection methodology studied was for UK general practice, which is designed to process several thousand applicants per year and targets six core attributes identified in a multi-method job-analysis study (empathy, communication skills, problem-solving, professional integrity, coping with pressure, and clinical expertise).[10,11] The selection process involves three stages:

1. long-listing eligibility checks;
2. short-listing via two machine-marked tests including (i) a clinical problem-solving test comprising questions that require applicants to apply clinical knowledge to solve problems reflecting diagnostic processes, or to develop management strategies for patients; and (ii) a situational judgement test targeting non-academic attributes (empathy, integrity, coping with pressure), where applicants are presented with written depictions of professional dilemmas they may encounter at work and are asked to identify an appropriate response from a list of alternatives;[12] and

F Patterson, BSc, MSc, PhD, AcSS, FRCGP (Hon), principal researcher, University of Cambridge and director, Work Psychology Group. **F Lievens**, PhD, professor, Department of Personnel Management and Work and Organizational Psychology, Ghent University, Belgium. **M Kerrin**, BSc, MSc, PhD, director, Work Psychology Group, Derbyshire. **N Munro**, MMEd, DPhil, FRCGP, FRCP(Edin), honorary senior clinical research fellow, Department of Health Care Management and Policy, University of Surrey. **B Irish**, BSc, MMEd, FRCGP, director of GP education, Health Education Southwest, Taunton.

How this fits in

Internationally there is limited research evidence available exploring the predictive validity of selection methods for entry into postgraduate medicine. This study builds on, and extends previous research by triangulating evidence from three longitudinal studies. The predictive validity of each selection method (a clinical problem-solving test, a situational judgement test and a selection centre) for various outcomes (including supervisor assessments 1 year into training and performance in end-of-training licensure exams) was examined. Results show that each of the selection methods is a significant independent predictor of trainee performance 1 year into training and for their end-of training competence in the licensure exams. The paper highlights the challenges of conducting predictive validation studies for any selection system (for example, restriction of range, defining appropriate outcome measures). Although there is clearly scope for improvements, compared to selection systems used in many other occupations, the UK GP selection system shows promising evidence for the predictive validity of the three-part methodology.

3. a selection centre using job-relevant simulations (patient consultation, group and written simulation exercises) to target both clinical and non-clinical attributes.[11,13]

Typically, 10–20% of applicants are rejected at short-listing, with a further 20–30% selected out at the final-stage selection centre. Initial evidence of the predictive validity of the selection system has been demonstrated at 3 months into training.[7]

This study substantially expands on a preliminary validation study to address the following three research questions:

1. Are the short-listing tests valid, that is, do they predict performance in the final stage of selection?
2. Does performance at selection predict subsequent job performance 1 year into training, as rated by supervisors?
3. Does performance at selection predict end-of-training competence (such as, performance in licensing examinations)?

METHOD

Design and sampling

Figure 1 provides a flowchart showing the study design and sampling at each time point, to address the research questions.

Sample 1: predictive validity of the short-listing tests. Selection data were collected during the 2007 annual recruitment round for UK GP training. Applicants meeting the long-listing eligibility requirements in the selection process completed both the clinical problem-solving test and the situational judgement test, for short-listing purposes (n = 8399). Scores from each test were equally weighted in determining short-listing outcomes, with successful applicants subsequently invited to the selection centre.
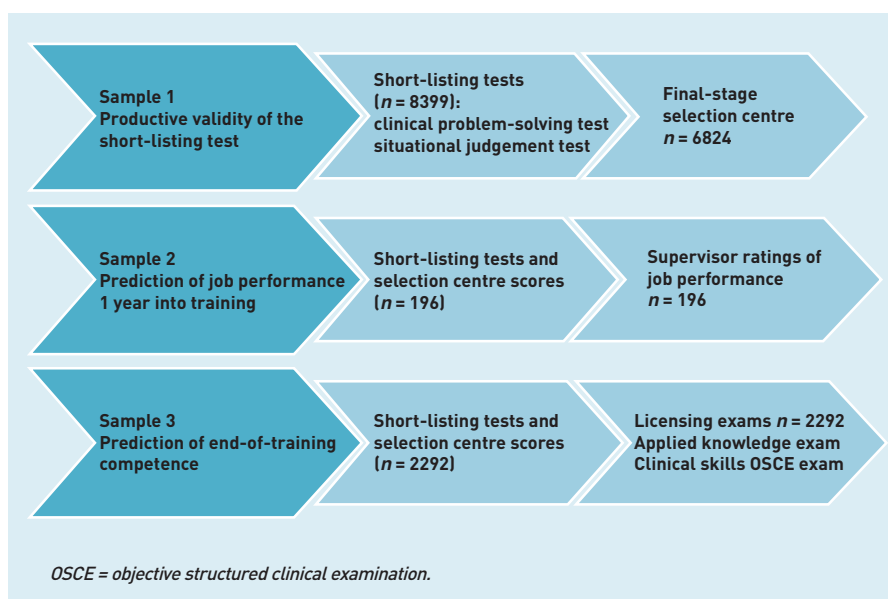

Figure 1. Flowchart of the three-part study showing the longitudinal research design.

OSCE = objective structured clinical examination.

Performance at the selection centre (total score across three job simulation exercises; $n = 6824$) was used as the outcome measure, as the primary aim of short-listing is to identify applicants who are likely to perform well at the selection centre. Initial evidence of the predictive validity of the clinical problem-solving test and a pilot version of the situational judgement test has been reported.[14]

Sample 2: prediction of job performance 1 year into training. A longitudinal design was used, which replicates the initial validation study[7] by tracking the performance of GP registrars 1 year into training. In 2008, a convenience sample of 490 GP registrars in five UK regions, who had completed a GP placement during their first year of training, were invited by their regional education and training director to participate in the study, resulting in 196 usable responses (response rate 40%). Note that accessing a sufficiently large sample is practically difficult, as a relatively low proportion of trainees entering training in general practice complete a GP placement in their first year of training (approximately 17%). Trainees in a GP placement were specifically targeted, rather than those entering training in a hospital environment, as the selection criteria identified from the original job analysis study were tailored specifically for general practice.[10] Clinical supervisors ($n = 196$) evaluated the registrars' performance on the six target attributes, using a 24-item inventory adapted from the original validation study. An example item for coping with pressure was 'is clear and rational when dealing with difficult issues or situations'. Ratings were made on a 6-point Likert-type scale (1 = needing significant development, to 6 = clearly demonstrated). Supervisors were blind to access to the GP registrars' selection scores when making their ratings. It was made clear that the data would be confidential and used for research purposes only.

Sample 3: Prediction of end-of-training competence. A retrospective longitudinal design was used to evaluate prediction of end-of-training competence in terms of licensing examination performance. GP registrars in the UK must complete a membership examination offered by the Royal College of General Practitioners (MRCGP) to practise independently after completion of training. The final MRCGP examination[15] comprises an applied knowledge examination with 200 items covering clinical medicine and evidence-based clinical practice, and a clinical skills objective structured clinical examination (OSCE), comprising 12 patient-simulated stations, designed to test clinical, professional and practical skills. Data from both examinations were collected for GP registrars during February 2008 to May 2009 (comprising six diets of the applied knowledge examination and five diets of the clinical skills OSCE) and were compared with their performance at selection.

Data analysis
All analyses were conducted using SPSS (version 15.0 for Windows). Pearson product–moment correlations and regression analyses were used to examine the predictive validity of the selection methods. In analysing associations between each short-listing test and selection centre scores, coefficients corrected for multivariate restriction of range are reported.[16] This is important, as uncorrected correlations that are computed on the basis of only the selected pool of applicants underestimate the size of correlations between variables. Therefore, corrected correlations are a more accurate reflection of the 'true' association between selection methods (predictors) and outcome variables.

A two-stage selection process was used. First, only candidates who passed the cut-off and top-down selection, determined on the basis of a composite of the knowledge test and situational judgement test, proceeded to the next stage. A second selection occurred because only candidates who passed the top-down selection, determined on the basis of a composite of the various dimensions of the selection centre, were selected. Given that selection was based on a composite on two occasions, the correlations were corrected for indirect range restriction,[17] using the multivariate range-restriction formulas of Ree *et al*.[18] Specifically, the two-stage approach delineated by Sackett and Yang was followed;[16] the selection centre group ($n = 196$) was treated as the restricted group and the short-listed group ($n = 6542$) as the unrestricted group, and then the multivariate range-restriction formulas were applied to the uncorrected correlations. Next, the short-listed group was treated as the restricted group and the initial applicant pool ($n = 8399$) as the unrestricted group, and the multivariate range-restriction formulas were applied to the corrected correlations. Statistical significance was determined prior to applying the corrections.[16]

Table 1. Sample demographics

	n	Mean age, years	Female, %	Male, %	White UK and Irish, %	Asian, %	Black, %	Other, %	UK trained,%	Non-UK trained, %
Sample 1 (all)	8399	31.1	52	48	33	46	8	12	51	49
Sample 1 (after short-listing)	6824	30.9	54	46	37	42	6	13	57	43
Sample 2	196	29.4	60	40	43	41	4	10	75	25
Sample 3	2292	30.1	56	43	49	35	3	11	73	27

Percentages do not all add up to 100, owing to missing demographic data.

RESULTS

All variables in the study show score distributions close to normality, with adequate levels of variability (Table 1).

Sample 1: predictive validity of the short-listing tests

Each selection method demonstrates high internal reliability (Cronbach's α; clinical problem-solving test α = 0.88, situational judgement test α = 0.88, selection centre α = 0.87). There was a significant positive correlation between scores on the clinical problem-solving test and situational judgement test (r = 0.53, P<0.001, n = 8399), suggesting that these selection methods have both common and independent variance (that is, a moderate overlap in what they are measuring but also some differences). The corrected correlation between scores on the clinical problem-solving test and selection centre is r = 0.47 (P<0.001), and between the situational judgement test and selection centre is r = 0.58 (P<0.001, n = 6824). This indicates that both short-listing tests are moderately strong predictors of selection centre performance. However, the best prediction of selection centre performance is the (equally weighted) combined score from both short-listing tests (r = 0.63, P<0.001). In comparing the predictive validity of the two short-listing tests, hierarchical regression analyses show that both tests offer incremental validity over the other in predicting selection centre performance. This analysis demonstrates the additional predictive value a selection method provides compared to other methods in the same study. The results show that the clinical problem-solving test predicted an additional 4% of the variance in selection centre scores over the situational judgement test, and the situational judgement test predicted an additional 11% of the variance in scores over the clinical problem-solving test (P<0.01).

Sample 2: prediction of job performance 1 year into training

The job performance inventory shows good internal reliability (Cronbach's α = 0.97). Table 2 presents the correlations between the selection methods (predictors) and supervisor ratings of job performance (outcome variable) after 1 year of training. All selection methods emerged as significant predictors of supervisor ratings of job performance (corrected r ranging from 0.50 to 0.56, P<0.001), indicating that each method has predictive validity. Multivariate regression analyses demonstrate that the strongest prediction of job performance ratings is a combination of all three selection methods (corrected r = 061, P<0.001). In comparing the predictive

Table 2. Correlations between selection methods and supervisor ratings of job performance after 1 year (sample 2), n = 196

			Correlation		
	Mean	SD	1	2	3
Selection methods (predictors)					
1. Clinical problem-solving test	78.88	9.02	–	–	–
2. Situational judgement test	637.87	34.31	0.50	–	–
3. Selection centre	3.32	0.39	0.30	0.43	–
Outcome variable					
4. Supervisor ratings	4.63	0.73	0.36 (0.54)	0.37 (0.56)	0.30 (0.50)

SD = standard deviation. Correlations between parentheses were corrected for multivariate range restriction.
Correlations ≥0.14 are significant at P<0.05; correlations ≥0.19 are significant at P<0.01.

Table 3. Correlations between the selection methods and end-of-training assessments (sample 3), *n* = 2292

	Mean	SD	Correlation			
			1	2	3	4
Selection methods (predictors)						
1. Clinical problem-solving test	80.08	8.14	–	–	–	–
2. Situational judgement test	640.13	31.66	0.40	–	–	–
3. Selection centre	3.34	0.36	0.24	0.32	–	–
Outcome variables						
4. End-of-training applied knowledge test	0.26	0.90	0.73 (0.85)	0.43 (0.69)	0.24 (0.41)	–
5. End-of-training clinical skills OSCE	0.20	0.80	0.38 (0.55)	0.43 (0.57)	0.32 (0.41)	0.41

OSCE = objective structured clinical examination. SD = standard deviation. Correlations between parentheses were corrected for multivariate range restriction. All correlations are significant at P<0.001.

validity of the three selection methods, hierarchical regression analysis shows that the situational judgement test and clinical problem-solving test explain a significant amount of additional variance in performance ratings over each other (6% and 4%, respectively). The selection centre demonstrates incremental validity in predicting job performance ratings, when compared to the combination of the situational judgement test and clinical problem-solving test (explaining an additional 2% of the variance), indicating that the selection centre adds value over the short-listing tests in this respect.

Sample 3: prediction of end-of-training competence (licensing exam)

Within this sample (*n* = 2292), results from all diets of the applied knowledge examination were standardised to enable overall analysis of the results, and the same was done for the clinical skills OSCE. The association between performance on the applied knowledge examination and clinical skills OSCE was r = 0.41, *P*<0.001 (using mean scores). Table 3 presents the correlations between the selection methods (predictors) and performance on the applied knowledge examination and clinical skills OSCE (outcome variables). Again, correlations corrected for multivariate range restriction are also provided.[16] All selection methods emerged as significant predictors for both examinations (corrected r ranging from 0.41 to 0.85, *P*<0.01), indicating that each method has good predictive validity.[19] The clinical problem-solving selection test showed a particularly strong correlation with the applied knowledge examination and this is likely to reflect the similarity in content between these two assessments. While the situational judgement test and

selection centre correlated similarly with both the applied knowledge examination and the clinical skills OSCE, correlations between the clinical problem-solving selection test and the clinical skills OSCE were substantially smaller than with the applied knowledge examination. In comparing the selection methods, hierarchical regression analyses were used to examine the incremental validity of each method in predicting performance in the two examinations. The situational judgement test and clinical problem-solving test each explain a significant amount of additional variance in the scores in the applied knowledge examination and clinical skills OSCE over each other (2% and 8% respectively for the situational judgement test, and 27% and 5% respectively for the clinical problem-solving test). The selection centre also demonstrated incremental validity (added value) over the situational judgement test and clinical problem-solving test combined, in predicting clinical skills OSCE scores (2% additional variance).

DISCUSSION
Summary
Results show both short-listing tests demonstrate good predictive validity with regard to performance on job simulations at the selection centre, with the strongest prediction offered by a combination of the clinical problem-solving test and situational judgement test. Although subject to issues regarding restriction of range, the results suggest that the short-listing process is effective in identifying applicants who go on to perform well at the final stage of selection. In relation to the second key research question, all three selection methods (both tests and the selection centre) were effective predictors of supervisor ratings of job performance in core attributes after 1 year of training. This finding, which builds upon the results of the authors' initial validation study,[7] provides encouraging evidence that the selection system is successful in identifying those candidates who will go on to become effective clinicians. Further, the high-fidelity selection centre job simulations provided additional incremental validity over the short-listing tests in predicting performance in interpersonally-oriented attributes such as empathy and communication skills, suggesting that this may represent the most valuable contribution of a selection centre over other methods. However, the sample size and response rate in this analysis was relatively small and so the results should be treated with some caution.

The final research question asks whether the selection system demonstrates predictive validity in relation to performance in end-of-training licensing examinations. The results indicate that all three selection methods were significant predictors of performance in both licensure examinations. The findings show that performance on the clinical problem-solving test is highly correlated with subsequent performance in an end-of-training knowledge examination targeting declarative (clinical) knowledge, which is to be expected.

The situational judgement test assesses a range of important (non-academic) professional attributes (empathy, integrity, and coping with pressure) that are important aspects of in-training job performance and end-of-training competence.[12] Compared to clinical knowledge tests, situational judgement tests target procedural knowledge and awareness of what are effective courses of action in a given situation, relating to important professional attributes.

The correlation between the selection centre and the clinical skills OSCE is moderately lower than for the two short-listing tests, although a corrected validity coefficient of $r = 0.41$ is substantial compared to many other selection tools.[20] In addition, it might be expected that the correlation between the selection centre and OSCE should be higher, as each assessment tool uses a similar test modality (such as, use of work samples and simulations). It can be argued that the level of correlation between the selection centre and OSCE is to be expected, since selection centres are based on a multi-trait multi-method approach,[20] and are designed to evaluate candidates' aptitude in relation to important non-clinical domains (for example, empathy and integrity). By contrast, an OSCE is a high-fidelity assessment of clinical competence, testing declarative knowledge in addition to important non-clinical domains.

Further, although selection centres are relatively expensive selection tools, the results show that the selection centre adds significant incremental validity over the two short-listing tests in predicting subsequent job performance (sample 2) and end-of-training competence (sample 3), especially for domains such as empathy and communication, which are crucial for those entering training in general practice.[10] When comparing all three selection methods, the best prediction of end-of-training competence is a combination of all three methods (combining both low- and high-fidelity assessments), as each adds something unique in predicting job performance and training outcomes.

Strengths and limitations
It could be argued that this study simply confirms that those applicants performing well at test-taking at the outset (selection) also perform well during tests at the end of training (licensure examinations). However, the validation evidence presented here also includes assessments of typical work performance (that is, independent supervisor in-training assessments of performance in practice) and that, in combination, this triangulation of results from this three-part study provides good evidence of the predictive validity of the selection methodology for entry into postgraduate training. A recognised weakness in all predictive validity studies is that correlations can only be computed for the selected pool of applicants, since some candidates are selected out in the process. However, uncorrected correlations tend to underestimate the size of correlations between variables and there are accepted methods for statistically correcting for this issue.

Comparison with existing literature
This study extends initial evidence regarding the validity of methods used for selection into UK general practice training,[7,14] by demonstrating the predictive validity of the selection system in three ways, including prediction of performance at the end of training. This is an important finding as this is the first study to explore the long-term predictive validity of postgraduate selection methods. Compared to large-scale meta-analytic studies of the predictive validity of selection methods for most other occupational groups, the size of validity coefficients for GP selection in the UK is substantial.[19,21]

Implications for research
A key question for future research relates to what is known as the 'criterion problem' in selection research: what outcomes are the selection methodology intending to predict specifically? For example, further evaluations could include an analysis of the emerging workplace-based assessment data, especially for trainees who are experiencing difficulties in training. However, executing validation studies is complex in practical terms, since researchers would rarely use one single predictor to make selection decisions and applicants will be judged on multiple selection criteria

Discuss this article
Contribute and read comments about this article on the Discussion Forum: **http://www.rcgp.org.uk/bjgp-discuss**

(depending on the stage in education and training pathway). A significant problem is with regard to accessing the appropriate criterion (outcome) data to validate a selection methodology. Often the criteria used to measure performance in the job role do not match the criteria used for selection. Conversely, sometimes the criterion and predictor are very similar (for example, using knowledge-based tests at selection to predict knowledge-based licensure examination performance), which may lead to problems of common method variance. Ideally, predictor scores should not be used to take selection decisions until after a predictive validation study has been conducted. Practically (and ethically), this is difficult to achieve and so statistically correcting for issues of restriction of range is an important consideration and a widely accepted approach when interpreting validity data in selection research.

The results presented here have implications for designing selection methodology in other medical and healthcare specialties internationally,[1,22–24] and for the design of selection methods for medical school admissions. The study would argue that the cornerstone to effective selection is in conducting a detailed job analysis study to accurately identify valid selection criteria. The original job analysis study has been repeated recently,[25] showing that the nature of the GP role is broader and potentially more complex than in the past and this is an important consideration for designing the selection system. The overall level of predictive validity could be improved further and additional systematic validation studies are required. In addition, the results presented here could be used to review the cost efficiency of the methodology, with an aim to optimise both efficiency and effectiveness in the future.[26] Previous research has tended to focus on evaluating the predictive validity of individual selection methods, whereas future research should focus on evaluating the optimal combination of methods using programmatic approaches to selection system design.

REFERENCES

1. Prideaux D, Roberts C, Eva K, *et al*. Assessment for selection for the health care professions and specialty training: international consensus statement and recommendations. *Med Teacher* 2011; **33(3):** 215–223.

2. Jefferis, T. Selection for specialist training. *BMJ* 2007; **334:** 1302–1304.

3. Patterson F, Ferguson E, Knight A. Selection for medical education and training. In: Swanick T (ed). *Understanding medical education: Evidence, theory and practice.* Chichester: Wiley-Blackwell, 2013: 352–365.

4. Patterson F, Ferguson E. Testing non-cognitive attributes in selection centres: How to avoid being reliably wrong? *Med Educ* 2012; **46(3):** 240–246.

5. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. *BMJ* 2002; **324(7343):** 952–957.

6. James D, Yates J, Nicolson S. Comparison of A level and UKCAT performance in students applying to UK medical and dental schools in 2006: cohort study. *BMJ* 2010; **340:** c478.

7. Patterson F, Ferguson E, Norfolk T, Lane P. A new selection system to recruit general practice registrars: preliminary findings from a validation study. *BMJ* 2005; **330:** 711–714.

8. Eva KW, Reiter HI, Trinh K, *et al*. Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ* 2009; **43(8):** 767–775.

9. de Oliveira GS Jr, Akikwala T, Kendall MC, *et al*.. Factors affecting admission to anesthesiology residency in the United States: choosing the future of our specialty. *Anesthesiology* 2012; **117(2):** 243–251.

10. Patterson F, Ferguson E, Lane P, *et al*. A competency model for general practice: implications for selection and development. *Br J Gen Pract* 2000; **50:** 188–193.

11. Irish B, Patterson F. Selecting general practice registrars — where next? *Br J Gen Pract* 2011; **60(580):** 849–852.

12. Patterson F, Ashworth V, Kerrin M, *et al*. Evaluating situational judgement tests to assess non-academic attributes in postgraduate selection. *Med Educ* 2012; **46(9):** 850–868.

13. Plint S, Patterson F. The way forward for recruitment into specialty training in the UK: Identifying critical success factors. *Postgrad Med* J 2010; **86:** 323–327.

14. Patterson F, Baron H, Carr V, *et al*. Evaluation of three short-listing methodologies for selection into postgraduate training in general practice. *Med Educ* 2009; **43:** 50–57.

15. Munro N, Hawthorne K, Denney M, *et al*. Developing a new clinical skills assessment for licensing UK general practitioners; the why and the how. *Educ Prim Care* 2006; **17:** 301–310.

16. Sackett P, Yang H. Correction for range restriction: an expanded typology. *J Appl Psychol* 2000; **85(1):** 112–118.

17. Thorndike RL. *Personnel selection: test and measurement techniques.* New York, NY: Wiley, 1949.

18. Ree MJ, Carretta TR, Earles JA, Albert W. Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulas. *J Appl Psychol* 1994; **79:** 298–301.

19. Schmidt F, Hunter J. The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychol Bull* 1998; **124(2):** 262–274.

20. Patterson F. Selection for medical education, training and practice. In: Walsh K (ed). *Oxford textbook of medical education.* Oxford: Oxford University Press, 2013, pp: 385-397.

21. Sackett P, Lievens F. Personnel selection. A*nnu Rev Psychol* 2008; 59: 16.1–16.32.

22. Patterson F, Carr V, Zibarras L, *et al*. New machine-marked tests for selection into core medical training: Evidence from two validation studies. *Clin Med* 2009; **9(5):** 1–4.

23. Patterson F, Ashworth V, Mehra S, Falcon H. Can situational judgement tests be used to select for dental foundation training? *Br Dent J* 2012; **213:** 23–26.

24. Patterson F, Lievens F, Kerrin M, *et al*. Designing selection systems for medicine: The importance of balancing predictive and political validity for high stakes selection. *International Journal of Selection and Assessment* 2012; 20(4): 486–496.

25. Patterson F, Tavabie A, Denney M, *et al*. A new competency model for general practice: Implications for selection, training and careers. *Br J Gen Pract* 2013; **63:** 249–250.

26. Irish B, Carr A, Sowden D, *et al*. Recruitment into specialty training in the UK. BMJ Careers 2011; 12 Jan. http://careers.bmj.com/careers/advice/view-article.html?id=20001789 (accessed 11 Oct 2013).