BMJ Open

Performance at medical school predicts success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-046615
Article Type:	Original research
Date Submitted by the Author:	04-Nov-2020
Complete List of Authors:	Ellis, Ricky; University of Aberdeen Institute of Applied Health Sciences, ; Nottingham University Hospitals NHS Trust, Urology Department Scrimgeour, Duncan; Aberdeen Royal Infirmary, Department of Colorectal Surgery Brennan, Peter ; Portsmouth Hospitals NHS Trust, Department of Oral and Maxillofacial Surgery Lee, Amanda; University of Aberdeen Institute of Applied Health Sciences Cleland, Jennifer; Lee Kong Chian School of Medicine
Keywords:	MEDICAL EDUCATION & TRAINING, SURGERY, Adult surgery < SURGERY

Performance at medical school predicts success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination

R Ellis[1,2,3], DSG Scrimgeour[1,3,4], PA Brennan[3], AJ Lee[5], J Cleland[6]

[1]Institute of Applied Health Sciences, University of Aberdeen and [2]Urology Department, Nottingham University Hospitals, Nottingham and [3]Intercollegiate Committee for Basic Surgical Examinations and [4]Department of Colorectal Surgery, Aberdeen Royal Infirmary, Aberdeen and [5]Department of Medical Statistics, University of Aberdeen, [6]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.

Correspondence to: Mr Ricky Ellis

Centre for Healthcare Education Research and Innovation,

University of Aberdeen, Aberdeen, AB25 2ZD

Rickyellis@nhs.net

Manuscript word count: 3,438

ABSTRACT

Background: Identifying predictors of success in post-graduate examinations can help guide the career choices of medical students and may aid early identification of trainees requiring extra support to progress in specialty training. We assessed whether performance on the Educational Performance Measurement (EPM) and Situational Judgement Test (SJT) used for selection into Foundation Training predicted success at the Membership of the Royal College of Surgeons (MRCS) examination.

Methods: This was a longitudinal, cohort study using data from the UK Medical Education Database (https://www.ukmed.ac.uk). UK medical graduates who had attempted Part A (n=1,975) and Part B (n=630) of the MRCS between 2013-2017 were included. Chi-squared and independent t-tests were used to examine the relationship between medical school performance and sociodemographic factors with success at MRCS Part A and B. Logistic regression was employed to identify independent predictors of MRCS performance.

Results: For every additional EPM decile point gained the chances of passing MRCS at first attempt increased by 52% for Part A (odds ratio 1.52 [95% confidence interval (CI) 1.46-1.60]) and 27% for Part B (1.27 [1.18-1.38]). For every point awarded for additional degrees in the EPM, candidates were 29% more likely to pass MRCS Part A first time (1.29 [1.12-1.48]).

SJT score was not a statistically significant independent predictor of MRCS Part A or Part B success (P=0.182 and P=0.125 respectively).

Conclusion: This is the first study to investigate the relationship between medical school and foundation training selection performance with performance at a high stakes UK postgraduate surgical examination. This study demonstrated the EPM's independent predictive power and found that medical school performance deciles are the most significant measure of predicting later success in the MRCS. These findings can be used by medical schools, training boards and workforce planners to inform evidence-based and contemporary selection and assessment strategies.

Key words: MRCS, UKFPO, situational judgement test, Medical School, Surgery, Postgraduate examinations

Strengths and limitations of this study

- The first study to investigate the relationship between medical school performance with performance at a high stakes UK postgraduate surgical examination.
- A large retrospective longitudinal cohort study using the UKMED database.

- Defines the relationship between each Foundation Program selection tool and success at the MRCS.
- Highlights group-level attainment differences between sociodemographic groups for further investigation.
- MRCS success at first-attempt used as a predictor of future performance in surgical training.

INTRODUCTION

Progression through the United Kingdom (UK) medical education and training pathway is based on performance on a series of index assessments, starting with examination performance prior to entry to medical school and (typically) ending with Royal College Fellowship examinations. Each assessment is designed to ensure appropriate standards for stage of training and to ultimately safeguard patients.(1,2)

Performance at each stage also has implications on career progression. In the UK, doctors with higher academic scores during medical school are more likely to be offered their first choice of UK Foundation Programme (UKFP) training post on graduating.(3) Those with higher academic scores during medical school are also more likely to be offered a training place in a more competitive specialty.(4)

Studies have already demonstrated the validity of academic performance during medical school in predicting performance during Foundation Training.(5-7) However, there is little research on the association between medical school performance and performance during specialty training in the UK. What research does exist was carried out before standardized markers of medical school performance were introduced (see later) and can therefore not be relied on in terms of illuminating contemporary patterns of performance.(8) This is an important gap in the literature as research from other contexts indicate that examination results taken during and shortly after medical school predict later performance on board certification examinations, and patient complaints.(9-14) Furthermore, if early assessments do not predict later performance, then their fitness for purpose as markers of performance and their use as gateways for progression in training are questionable.

At the time of this study the UK did not have a national licensing examination for graduating doctors. Instead, performance during medical school is measured within schools by the Educational Performance Measure, or EPM.(15) The EPM is calculated out of 50 points and comprises three parts (Table 1); medical school performance decile (points are awarded depending on a student's final EPM decile; ranging from 34 points for the 10th (lowest) decile to 43 points for students in the 1st (highest) decile); additional degrees, 0-5; and publications, 0-2. The EPM is an example of a programmatic assessment that grades satisfactory performance judged over time and by multiple assessments of several modalities.(16) The selection process for the UKFP couples the EPM with a situational judgement test, or SJT,(17-21) also scored out of 50 points, which tests the behaviours and attitudes expected of doctors as described in the General Medical Council's (GMC) Good Medical Practice.(2) The graduate's combined EPM plus

SJT score out of 100 is their application score for the UKFP.(22)

Table 1. The components of the Educational Performance Measure used to quantify performance at medical school. The EPM is combined with a candidates SJT score to create their total UKFP selection score.

Educational Performance Measure (EPM)	**Points available**
Medical School Performance Decile	34-43
Additional Degrees	0-5
Publications	0-2
Maximum combined EPM score	**50**

Situational Judgment Test (SJT)	**Points available**
Test Score	50

United Kingdom Foundation Program (UKFP) Selection Score	**Points available**
EPM	50
SJT	50
Maximum UKFP Selection Score	**100**

Our aim was to assess whether performance in medical school, EPM and SJT scores, could predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination. The MRCS examination is often taken by UK trainees during Foundation and Core surgical training years and comprises of two parts: Part A, a written examination with two papers and Part B, an objective structured clinical examination (OSCE).(23) The MRCS is a high-stakes postgraduate assessment that is used as a gateway for applications for higher surgical training and is itself a good predictor of future surgical training outcomes.(24-26)

Given success in medical education and training is related to socio-demographic factors as well as academic ability,(3,27–29) we also examined the relationship between MRCS success and sociodemographic factors. These included graduate status on entry to medical school, gender and ethnicity. This analysis is timely given policy drivers in the UK to ensure that medical school and postgraduate assessments are fair,(1) and the pending imposition of a once-off high-stakes test, the Medical Licensing Assessment (MLA).(30)

Use of linked individual-level data from the UK Medical Education Database (UKMED: https://www.ukmed.ac.uk/) enabled a national-level analysis, drawing on data from sources including medical school assessment, Foundation Programme selection and postgraduate assessment outcomes.(31)

METHODS

A longitudinal retrospective cohort study was conducted on UK medical graduates who had attempted either the Part A (written) or the Part B (clinical) MRCS examination from September 2013 to May 2017.

The UK Medical Education Database (UKMED: https://www.ukmed.ac.uk/) was used to access linked data from UK medical schools and the four Royal Colleges of Surgeons in the UK and Ireland. All counts have been rounded according to Higher Education Statistics Agency (HESA) data standards to ensure person-level anonymity.(32)

The following data were extracted: self-declared gender, graduation status at the time of entry to medical school and self-reported ethnicity demographics, medical school Educational Performance Measure decile, additional degree and EPM publication scores, SJT score and MRCS Part A and B first attempt result. Candidate first attempt results were used as they have been shown to be the best predictor of future performance in postgraduate examinations.(33)

Except for SJT and EPM scores, all variables were subsequently dichotomized. Graduation status was defined as “yes” if candidates had obtained a degree prior to entering medicine. Self-declared ethnicity was coded as “white” or “non-white” as used in similar studies to enable powered analysis of smaller cohorts.(24,25) Part A and B MRCS performance was categorized as "pass"'or "fail" at first attempt.

Statistical analysis

All analyses were conducted using SPSS® v22.0 (IBM, Armonk, New York, USA). A chi-squared test was initially employed to determine any associations with first attempt MRCS pass/fail outcomes. The relationship between SJT, EPM decile, additional degrees, EPM publication scores and Part A and Part B MRCS first attempt success was examined using independent t-tests since the distribution of scores was normal. Pearson's correlation coefficients were calculated for Foundation Programme selection scores and the MRCS Part A (the written component of the MRCS examination).

Logistic regression models were developed to identify potential independent predictors of first attempt success at Part A and B MRCS. Any variable with $P<0.10$ on univariate analysis was entered into the logistic regression model. All potential predictors with $P>0.05$ in the full model were subsequently removed until only statistically significant predictors remained in the final model. Potential interactions between the remaining significant predictors were also examined.

The Intercollegiate Committee for Basic Surgical Examinations (ICBSE) and its Internal Quality Assurance Subcommittee, which monitors standards and quality, approved this study. The highest standards of security, governance and confidentiality were ensured when storing, handling and analysing identifiable data.

RESULTS

Part A MRCS

A total of 1,975 UK medical graduates attempted Part A MRCS between September 2013 and May 2017. 55% (n=1085) passed Part A MRCS at their first attempt. 66% of candidates (n=1300), were male, 58% were white (n=1125), and 76% had not undertaken a prior degree before entering medicine (n=1490). Mean (standard deviation [SD]) total EPM and SJT scores for candidates who had attempted Part A MRCS were 42.9 (3.33) and 39.8 (3.39) respectively.

Pass rates for Part A MRCS by gender, graduate on entry to medicine status and ethnicity are shown in Table 2. Differences in pass rates were statistically significant for: gender (59.4% males *vs.* 46.7% females, $P<0.001$), graduate status (58.6% no prior degree *vs.* 44.1% prior degree, $P<0.001$) and ethnicity (59.8% white *vs.* 49.2% non-white, $P<0.001$).

Table 2. MRCS first attempt pass rates by gender, ethnicity and graduation status for UK medical graduates

Variable	Part A (n = 1975)	Part B (n = 630)
Graduate on entry to medicine		
No	58.6% (875/1490)	80.0% (405/510)
Yes	44.1% (210/480)	68.6% (85/120)
Missing	n=0	n=0
p-value	< 0.001	0.007
Gender		
Male	59.4% (770/1300)	77.5% (335/430)
Female	46.7% (315/675)	78.4% (155/200)
Missing	n=0	n=0
p-value	< 0.001	0.837
Ethnicity		
White	59.8% (675/1125)	84.5% (300/355)
Non-white	49.2% (410/830)	69.5% (190/275)
Missing	n=20	n=0
p-value	< 0.001	< 0.001

Note. All p-values presented are from chi-squared analysis. MRCS, Membership of the Royal College of Surgeons.

Univariate analysis of EPM and SJT scores are shown in Table 3. Candidates who passed Part A MRCS at first attempt had performed better in their SJT (40.3 [3.1] *vs.* 39.2 [3.6], P<0.001) and had scored higher for their total EPM (44.3 [2.9] *vs.* 41.2 [3.0], P<0.001) compared to those who failed at first attempt.

Table 3. Univariate analysis of EPM scores, SJT scores and MRCS Part A and Part B first attempt success.

Variable		MRCS Part A			MRCS Part B		
		Pass *n*=1085	**Fail** *n*=885	***P***	**Pass** *n*=490	**Fail** *n*=140	***P***
EPM Decile	Mean	40.2	37.5	<0.001	39.9	38.2	<0.001
	S.D.	2.4	2.5		2.5	2.6	
EPM Degree score	Mean	3.4	3.1	<0.001	3.4	3.2	0.027
	S.D.	0.7	0.9		0.7	0.8	
EPM Publications	Mean	0.8	0.6	<0.001	1.0	0.9	0.388
	S.D.	0.9	0.8		1.0	0.8	
SJT	Mean	40.3	39.2	<0.001	40.5	39.3	0.004
	S.D.	3.1	3.6		3.1	4.3	

Note. MRCS, Membership of the Royal College of Surgeons. EPM, Educational Performance Measure. SJT, Situational Judgment Test, N, number of candidates. S.D., Standard Deviation. *P*, *P* Value.

Table 4 shows Pearson's correlation coefficients between each Foundation Programme selection score and MRCS Part A. According to Cohen's guidelines (34) EPM degree score, EPM publication score and the SJT show statistically significant weak positive correlation with Part A scores. Whilst Total EPM and EPM Decile show statistically significant moderate correlation with MRCS Part A.

Table 4. Pearson's correlation coefficients between Foundation Programme selection scores and MRCS Part A Scores.

	Pearson Correlation	P Value
Total EPM	0.55	<0.001
EPM Decile	0.57	<0.001
EPM Degree Score	0.22	<0.001
EPM Publications	0.13	<0.001
SJT	0.23	<0.001

Table 5 shows the odds ratios (OR) and 95% confidence intervals (CI) for independent predictors of passing Part A MRCS at first attempt. For every additional EPM decile, the odds of passing Part A MRCS at first attempt increased by 52% (OR 1.52, 95% CI 1.46 to 1.60) and for every additional EPM degree point awarded, the odds of passing Part A on first attempt increased by 29% (OR 1.29, 95% CI 1.12 to 1.48). Neither EPM publication score or SJT score were independent predictors of Part A first attempt success (P=0.182 and P=0.222 respectively). MRCS candidates who entered medical school without a prior degree were more than twice as likely to pass Part A compared to those who entered medical school as graduates (OR 2.45, 95% CI 1.91 to 3.15). There was a statistically significant interaction between ethnicity and gender in the final Part A MRCS regression model with white males more likely to pass, P=0.001.

Table 5. Predictors of pass at first attempt at Part A (n=1975) and Part B (n=630) MRCS for UK medical graduates.

	Part A	**Part B**

Variable	Odds ratio (95% CI)	Odds Ratio (95% CI)
Educational Performance Measure Decile	1.52 (1.46 to 1.60)	1.27 (1.18 to 1.38)
Educational Performance Measure Additional Degree points	1.29 (1.12 to 1.48)	-
Graduate on entry into medicine Non-Graduates vs. Graduates	2.45 (1.91 to 3.15)	2.56 (1.57 to 4.15)
Gender* Males vs. Females	1.37 (0.98 to 1.93)	-
Ethnicity* White vs. Non-White	0.75 (0.52 to 1.08)	2.22 (1.47 to 3.36)
Ethnicity*Gender	2.10 (1.34 to 3.28)	-

MRCS, Membership of the Royal College of Surgeons; CI, Confidence Interval. *P=0.001 for interaction between Ethnicity and Gender in MRCS Part A Model.

Part B MRCS

In total, 630 UK medical graduates attempted Part B of the MRCS from September 2013 to May 2017. 77.8% (n=490) of candidates passed Part B MRCS at first attempt. Unsurprisingly the demographics for Part B MRCS were similar to those observed for Part A MRCS candidates; 68% of candidates were male (n=430), 56% were white (n=355), and 81% had not undertaken a previous degree (n=510). The mean (SD) total EPM and SJT scores for UK graduates who had attempted Part B MRCS were 43.8 (3.16) and 40.2 (3.45) respectively.

Pass rates for Part B MRCS by gender, graduate on entry to medicine status and ethnicity are shown in Table 2. There was no significant difference in Part B MRCS first attempt pass rates between males and females (77.5% *vs.* 78.4% respectively P=0.837). Differences in pass rates were statistically significant for graduate status (80% no prior degree *vs.* 68.6% prior degree, P=0.007) and ethnicity (84.5% white *vs.* 69.5% non-white, P<0.001).

Univariate analysis of EPM and SJT scores are displayed in Table 3. Those who passed Part B MRCS at first attempt had performed better in their SJT compared to those who failed at first attempt (40.5 [3.1] *vs.* 39.3 [4.3], P<0.001). Similarly, candidates who passed Part B at first attempt had scored higher in their total EPM (44.2 [3.1] *vs.* 42.3 [3.0], P<0.001).

Table 5 shows the logistic regression model for independent predictors of Part B MRCS first attempt. For every additional EPM decile the chances of passing Part B MRCS at first attempt increased by 27% (OR 1.27, 95% CI 1.18 to 1.38). SJT score was not found to be an independent predictor of Part B first attempt success (P=0.125) and unlike Part A, EPM degree points were not found to be a statistically significant independent predictor of Part B success (P=0.072).

Candidates who had not undertaken a previous degree before entering medicine were more than twice as likely to pass Part B MRCS compared to those who had undertaken a prior degree (OR 2.56, 95% CI 1.57 to 4.15). White UK medical graduates were over twice as likely to pass Part B MRCS at first attempt compared to non-white candidates (OR 2.22, 95% CI 1.47 to 3.36). There were no statistically significant interactions between any of the Part B MRCS variables in the final regression model.

DISCUSSION

We assessed the predictive validity of the Foundation Programme (FP) selection measures, the SJT and EPM, against the MRCS examination which is known to be a good predictor of future surgical training outcomes.(24-26) We found that EPM deciles independently predicted success at both Part A (written) and Part B (OSCE) of the MRCS. For every incremental EPM decile, candidates were significantly more likely to pass both Part A and Part B of the MRCS.

Points awarded in the EPM for additional degrees independently predict success in MRCS Part A but not Part B. Whilst points awarded for additional publications were not independent predictors of success in either exam, they do show a statistically significant but weak positive correlation with Part A scores. They will undoubtedly play a role in increasing the spread of applicant scores when combined with EPM decile and SJT points, creating a valuable tool for the Foundation Programme/UKFP when ranking thousands of applicants each year.

Our results add to previous studies which found that the EPM predicts performance during the FP,(6,7) and provides assurance that UK medical school assessments appropriately gauge student competence and readiness for practice. How the EPM, a local-level assessment, will sit alongside the proposed MLA, a national-level assessment, remains to be seen. Moreover, the predictive and incremental validity of the new MLA must be scrutinised in order to justify its financial cost and its burden on both students and the medical education system.

Our findings also align with the "academic backbone" concept; an idea that in medical education, current learning and achievement is dependent upon achievement at earlier stages.(8) This can be summarised simply as: medical students who are high achievers remain high achievers. Those who perform best in MRCS are more likely to achieve a specialty training (ST3) post at national selection, are more likely to progress through training with satisfactory Annual Review of

Competence Progression (ARCP) outcomes and are more likely to succeed at first attempt in their surgical fellowship (FRCS) examinations.(24–26) The nature of the outcome measures have changed since McManus et al.'s seminal study in 2013,(8) but the principles have not: the road to success for those who wish to pursue a successful career in surgery begins early.

Whilst candidates who passed both parts A and B of the MRCS on first attempt scored higher in their SJT than candidates who failed, and there was a statistically significant weak positive correlation with Part A scores, the SJT did not independently predict MRCS success after adjusting for EPM score. It is important to consider this finding in relation to the premise behind SJTs. SJTs are well-validated tools for large scale candidate selection processes.(17–21) They are designed to measure the expression of personality traits in hypothetical situations designed on the basis of what is expected in the job for which the individual is being assessed.(35,36) They encompass measurement of personal choice (e.g., what would be the best way to respond in this particular situation?).(37,38)

The Foundation Programme (FP) SJT is based on a job analysis of being a Foundation Doctor.(7) Significant correlation between SJT and EPM scores between schools has been identified.(39) Additionally both SJT and EPM scores are independently associated with the odds of successful completion of the FP, and SJT score offers a degree of incremental predictive validity over that provided by the EPM deciles, suggesting that it is capturing additional, relevant, information on applicants, as intended.(6) In short, research suggests that the FP SJT does what it was designed to do as well as succeeding in increasing the spread of candidates being ranked for foundation training posts (the arguments as to how it should be weighed in the FP selection process are outside the scope of this paper but we direct readers with an interest to other studies).(6,40,41) It was not designed to select for specialty training: where specialty training programmes use SJTs for selection, these have been designed specifically against the role of a trainee/resident in that speciality.(42-44) Given this, in retrospect it is unsurprising that the FP SJT does not independently predict performance on a post-graduate examination that tests the clinical knowledge, skills and professional attitudes expected of surgical trainees.

Attainment Differences

Interestingly, whilst candidates with no additional degree on application to the foundation program will receive a lower total EPM score, those who study medicine as an undergraduate degree (the norm in the UK) are more than twice as likely to pass MRCS Parts A and B on their first attempt. These results support previous studies that have found that students who enter medical school as graduates perform at least as well as undergraduate students in medical school examinations,(28,45) but the opposite pattern is seen in post-graduate training.(24,46,47) For the most part, older candidates (defined as 29 years or older in previous studies(23)) are assumed to be graduate-entrants, and this group are more likely to struggle during specialty training.(24,46,47) The reasons for this are unclear but may be related to competing time demands.(47)

However, it is important to note that the points awarded for additional degrees within the EPM could be because of a degree prior to entering medical school (that is, entering medicine as a graduate), or because of taking a year out during medical school to obtain a degree (intercalating). Given the existing literature on graduate-entrant performance later in training and the general literature on how students perform after intercalating,(48) we suggest that those who perform well on the EPM and the MRCS are likely to have intercalated rather than be graduate entrants to medicine. Future studies may wish to take a more forensic analysis of the associations between prior degrees (graduate-entry) and intercalated degrees and later performance to test this hypothesis.

As per previous studies,(23,24) a higher percentage of men passed Part A MRCS (written) on first attempt than women, although there was no significant difference by gender for the Part B clinical examination. Additionally, more white candidates passed both Parts A and B MRCS on first attempt compared to non-white candidates and ethnicity was found to be a significant independent predictor of Part B success in the logistic regression model. The interaction term between ethnicity and gender was a significant predictor of passing Part A, implying that white males are considerably more likely to pass MRCS Part A on first attempt.

Group differences by ethnicity and gender are well documented in other postgraduate medical examinations, including the Fellowship of the Royal College of Surgeons (FRCS) and the United States Medical Licensing Examination (USMLE).(24,49-51) The underlying reasons for differential attainment are likely to be complex and multi-factorial. Other studies have identified examiner bias in clinical examinations; examination questions can be inherently biased; and/or there may be actual differences between groups.(23,49,50,52) Analytical approaches using item analysis should be used to see if certain, unfair questions explain group-level differences in performance.(53) More broadly, work is required to examine differences between genders and different ethnic groups in terms of exposures, experiences and outcomes in medical training. It is imperative that such attainment gaps are explored if we as a profession are to ensure the complete eradication of structural inequalities between socio-demographic groups in order to improve fairness, diversity and representation within the workforce.

Strengths and Weaknesses

The strength of the current study is that it is one of the first to use the UKMED to examine the associations between sociodemographic factors, medical school performance and FP SJT outcomes on success at a high-stakes postgraduate surgical examination. The UKMED enabled a nationwide, multi-cohort analysis and our breakdown of the FP selection process into EPM and SJT allowed us to look separately at academic attainments and other factors.

There are some limitations of the study. Firstly, although candidates can take Parts A and B MRCS on multiple occasions, we used candidate first attempt results as the best predictor of future performance.(33) We often used the relatively blunt outcome measure of pass/fail as this is what is meaningful to those sitting the MRCS, and has been used in previous studies looking at factors which predict performance in the MRCS.(24) Secondly, when compared to previous studies,(23,33,49,52,54) we did not examine the relationship between first language and MRCS performance. Self-declared ethnicity data were combined into two discrete categories enabling powered analysis of smaller cohorts. Larger cohort sizes would enable a more granular analysis in order to minimise generalisations.(55) Finally, this analysis was based on retrospective quantitative data. A prospective study would have allowed us to examine more variables related to progression and differential attainment in surgical training. For example, being good at passing exams is linked to academic ability, but the wider education literature makes clear that non-cognitive factors such as motivation, time management and resilience are also relevant to performance.(56,57) If appropriate measures could be identified,(58) it would be interesting to compare graduates and MRCS candidates on these as well as additional sociodemographic factors, such as type of (high) school, medical school attended and social class.

CONCLUSION

Success at first attempt of MRCS Part A and B can be predicted from medical school performance (EPM decile score) but not from the Foundation Programme SJT score. Put simply, medical students who do well in terms of medical school examination performance and additional educational achievements remain strong performers later on in their careers. These results may help to guide career choices for students.

ABBREVIATIONS

ARCP: Annual Review of Competence Progression

EPM: Educational Performance Measure

FRCS: Fellowship of the Royal College of Surgeons Examinations

GMC: General Medical Council

HESA: Higher Education Statistics Agency

MLA: United Kingdom Medical Licensing Examination

MRCS: Intercollegiate Membership of the Royal College of Surgeons Examinations

SJT: Situational Judgement Test

UKFP: United Kingdom Foundation Program

USMLE: United States Medical Licensing Examinations

ACKNOWLEDGEMENTS

The authors would like to acknowledge Iain Targett at the Royal College of Surgeons of England, for his help with data collection and Gregory Ayre from the Intercollegiate Committee for Basic Surgical Examinations for their support during this project. Our thanks to members of the UKMED Research Group who provided useful feedback on an earlier version of this manuscript, and whose comments were helpful in refining the paper. The authors would also like to acknowledge Daniel Smith for his help with the UKMED database. Data Source: UK Medical Education Database ("UKMED"). UKMEDP043 extract generated on 25/07/2018. We are grateful to UKMED for the use of these data. However, UKMED bears no responsibility for their analysis or interpretation. The data includes information derived from that collected by the Higher Education Statistics Agency Limited ("HESA") and provided to the GMC ("HESA Data"). Source: HESA Student Records 2007/2008 to 2015/2016. Copyright Higher Education Statistics Agency Limited. The Higher Education Statistics Agency Limited makes no warranty as to the accuracy of the HESA Data, cannot accept responsibility for any inferences or conclusions derived by third parties from data or other Information supplied by it.

FUNDING

Royal College of Surgeons of England.

Royal College of Surgeons of Edinburgh.

Royal College of Surgeons of Ireland.

Royal College of Physicians and Surgeons of Glasgow.

CONFLICTS OF INTEREST

None to declare.

AUTHOR CONTRIBUTIONS

RE and DS wrote the first draft of the manuscript. RE and DS performed statistical analyses with AL's supervision. All authors reviewed and edited the manuscript. JC led the study proposal for access to UKMED data. All authors approved final draft of the manuscript.

REFERENCES

1. Promoting excellence: standards for medical education and training [Internet]. General Medical Council; 2015 [cited 2020 May 21]. Available from: www.gmc-uk.org/education/standards.asp

2. Good Medical Practice [Internet]. General Medical Council; 2013 [cited 2020 May 19]. Available from: www.gmc-uk.org/guidance.

3. Kumwenda B, Cleland JA, Prescott GJ, et al. Relationship between sociodemographic factors and selection into UK postgraduate medical training programmes: a national cohort study. BMJ Open. 2018 Jun;8(6):e021329.

4. Kumwenda B, Cleland J, Prescott G, et al. Relationship between sociodemographic factors and specialty destination of UK trainee doctors: a national cohort study. BMJ Open. 2019 Mar;9(3):e026961.

5. Patterson F, Ashworth V, Murray H, et al. Analysis of the Situational Judgement Test for Selection to the Foundation Programme 2013. Technical Report. [Internet]. Work Psychology Group; 2013 [cited 2020 May 19]. Available from: https://isfporguk.files.wordpress.com/2017/04/fy1-sjt-2013-technical-report.pdf

6. Smith DT, Tiffin PA. Evaluating the validity of the selection measures used for the UK's foundation medical training programme: a national cohort study. BMJ Open. 2018 Jul 1;8(7):e021918.

7. Medical Schools Council on behalf of the cross-stakeholder Project Group. Improving Selection to the Foundation Programme Final Report of the Parallel Recruitment Exercise. [Internet]. 2011 [cited 2020 Apr 28]. Available from: https://isfporguk.files.wordpress.com/2017/04/isfp-final-report-appendices.pdf

8. McManus I, Woolf K, Dacre J, et al. 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 Dec;11(1):242.

9. de Virgilio C, Yaghoubian A, Kaji A, et al. Predicting performance on the American Board of Surgery qualifying and certifying examinations: a multi-institutional study. Arch Surg Chic Ill 1960. 2010 Sep;145(9):852–6.

10. Tamblyn R, Abrahamowicz M, Brailovsky C, et al. Association between licensing examination scores and resource use and quality of care in primary care practice. JAMA. 1998 Sep 16;280(11):989–96.

11. Tamblyn R, Abrahamowicz M, Dauphinee D, et al. Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. JAMA. 2007 Sep 5;298(9):993–1001.

12. McCaskill QE, Kirk JJ, Barata DM, et al. USMLE step 1 scores as a significant predictor of future board passage in pediatrics. Ambul Pediatr Off J Ambul Pediatr Assoc. 2007 Apr;7(2):192–5.

13. Wenghofer E, Klass D, Abrahamowicz M, et al. Doctor scores on national qualifying examinations predict quality of care in future practice. Med Educ. 2009 Dec;43(12):1166–73.

14. Norcini JJ, Boulet JR, Opalek A, et al. The relationship between licensing examination performance and the outcomes of care by international medical school graduates. Acad Med J Assoc Am Med Coll. 2014 Aug;89(8):1157–62.

15. Educational Performance Measure (EPM) 2020 Framework [Internet]. UK Foundation Programme; 2020 [cited 2020 May 19]. Available from: https://foundationprogramme.nhs.uk/resources/

16. Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011;33(6):478–85.

17. Patterson F, Driver R. Situational Judgement Tests (SJTs). In: Patterson F, Zibarras L, editors. Selection and Recruitment in the Healthcare Professions [Internet]. Springer International Publishing; 2018 [cited 2020 Apr 27]. p. 79–112. Available from: http://link.springer.com/10.1007/978-3-319-94971-0_4

18. Patterson F, Knight A, Dowell J, et al. How effective are selection methods in medical education? A systematic review. Med Educ. 2016 Jan;50(1):36–60.

19. Mcdaniel M, Morgeson F, Finnegan E, et al. Use of Situational Judgment Tests to Predict Job Performance. J Appl Psychol. 2001 Sep 1;86:730–40.

20. McDaniel MA, Hartman NS, Whetzel DL, et al. Situational Judgment Tests, Response instructions and Validity: A Meta-Analysis. Pers Psychol. 2007 Mar;60(1):63–91.

21. Lievens F, Peeters H, Schollaert E. Situational judgment tests: a review of recent research. Pers Rev. 2008 Jun 6;37(4):426–41.

22. UK Foundation Programme 2020 Applicants Handbook. [Internet]. 2020 [cited 2020 May 19]. Available from: https://www.foundationprogramme.nhs.uk/wp-content/uploads/sites/2/2019/10/UKFP-2020-Applicants-Handbook-.pdf

23. Scrimgeour DSG, Cleland J, Lee AJ, et al. Which factors predict success in the mandatory UK postgraduate surgical exam: The Intercollegiate Membership of the Royal College of Surgeons (MRCS)? Surg J R Coll Surg Edinb Irel. 2018 Aug;16(4):220–6.

24. Scrimgeour DSG, Cleland J, Lee AJ, et al. Prediction of success at UK Specialty Board Examinations using the mandatory postgraduate UK surgical examination. BJS Open. 2019 Dec;3(6):865–71.

25. Scrimgeour D, Brennan P, Griffiths G, et al. Does the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination predict 'on-the-job' performance during UK higher specialty surgical training? Ann R Coll Surg Engl. 2018 Nov;100(8):669–75.

26. Scrimgeour DSG, Cleland J, Lee AJ, et al. Impact of performance in a mandatory postgraduate surgical examination on selection into specialty training: BJS Open. 2017 Oct;1(3):67–74.

27. Richens D, Graham TR, James J, et al. Racial and Gender Influences on Pass Rates for the UK and Ireland Specialty Board Examinations. J Surg Educ. 2016 Jan 1;73(1):143–50.

28. Puddey IB, Mercer A, Carr SE. Relative progress and academic performance of graduate vs undergraduate entrants to an Australian medical school. BMC Med Educ. 2019 May 22;19(1):159.

29. Woolf K, Cave J, Greenhalgh T, et al. Ethnic stereotypes and the underachievement of UK medical students from ethnic minorities: qualitative study. BMJ. 2008;337(a1220).

30. Medical Licensing Assessment [Internet]. General Medical Council; 2020 [cited 2020 May 19]. Available from: https://www.gmc-uk.org/education/medical-licensing-assessment

31. Dowell J, Cleland J, Fitzpatrick S, et al. The UK medical education database (UKMED) what is it? Why and how might you use it? BMC Med Educ. 2018 Dec;18(1):6.

32. Rounding and suppression to anonymise statistics | HESA [Internet]. [cited 2020 Aug 24]. Available from: https://www.hesa.ac.uk/about/regulation/data-protection/rounding-and-suppression-anonymise-statistics

33. McManus I, Ludka K. Resitting a high-stakes postgraduate medical examination on multiple occasions: nonlinear multilevel modelling of performance in the MRCP(UK) examinations. BMC Med. 2012 Jun 14;10(1):60.

34. Hemphill JF. Interpreting the magnitudes of correlation coefficients. Am Psychol. 2003;58(1):78–9.

35. Motowidlo SJ, Hooper AC, Jackson HL. Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. J Appl Psychol. 2006 Jul;91(4):749–61.

36. Patterson F, Ashworth V, Zibarras L, et al. Evaluations of situational judgement tests to assess non-academic attributes in selection. Med Educ. 2012 Sep;46(9):850–68.

37. Motowidlo SJ, Beier ME. Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. J Appl Psychol. 2010 Mar;95(2):321–33.

38. Patterson F, Kerrin M, Edwards H, et al. Validation of the F1 Selection Tools [Internet]. Work Psychology Group on behalf of Health Education England; 2015 [cited 2020 May 19]. Available from: https://isfporguk.files.wordpress.com/2018/06/validation_of_the_f1_selection_tools_report_final_for_publication-1.pdf

39. McManus IC, Harborne AC, Horsfall HL, et al. Exploring UK medical school differences: the MedDifs study of selection, teaching, student and F1 perceptions, postgraduate outcomes and fitness to practise. BMC Med. 2020 May 14;18(1):136.

40. Najim M, Rabee R, Sherwani Y, et al. The situational judgement test: a student's worst nightmare. Adv Med Educ Pract. 2015 Oct 5;6:577–8.

41. Mclachlan JC, Illing J. An Independant review of events relating to the use of Situational Judgement Tests in the selection for Foundation process 2013. Commissioned by Health Education England [Internet]. 2013 [cited 2020 May 19]. Available from: https://www.medschools.ac.uk/media/1932/hee-independent-review-of-the-sjt.pdf

42. Patterson F, Rowett E, Hale R, et al. The predictive validity of a situational judgement test and multiple-mini interview for entry into postgraduate training in Australia. BMC Med Educ. 2016 Dec;16(1):87.

43. Applicant Guidance. General Practice: Multi-Specialty Recruitment Assessment (MSRA) [Internet]. Health Education England; [cited 2020 May 19]. Available from: https://gprecruitment.hee.nhs.uk/Resource-Bank/Recruitment-Documents-Forms

44. Multi-Specialty Recruitment Assessment (MSRA) – Test Blueprint & Information [Internet]. Work Psychology Group on behalf of Health Education England; 2019 [cited 2020 May 19]. Available from: https://gprecruitment.hee.nhs.uk/Resource-Bank/Recruitment-Documents-Forms

45. Manning G, Garrud P. Comparative attainment of 5-year undergraduate and 4-year graduate entry medical students moving into foundation training. BMC Med Educ. 2009 Dec 22;9(1):76.

46. Reports on the progress of doctors in training split by postgraduate body. General Medical Council. [Internet]. [cited 2020 May 19]. Available from: https://www.gmc-uk.org/education/reports-and-reviews/progression-reports

47. Pyne Y, Ben-Shlomo Y. Older doctors and progression through specialty training in the UK: a cohort analysis of General Medical Council data. BMJ Open. 2015 Feb 1;5(2):e005658.

48. Jones M, Hutt P, Eastwood S, et al. Impact of an intercalated BSc on medical student performance and careers: a BEME systematic review: BEME Guide No. 28. Med Teach. 2013 Oct;35(10):e1493-1510.

49. Dewhurst NG, McManus C, Mollon J, et al. Performance in the MRCP(UK) Examination 2003–4: analysis of pass rates of UK graduates in relation to self-declared ethnicity and gender. BMC Med. 2007 May 3;5:8.

50. Rubright JD, Jodoin M, Barone MA. Examining Demographics, Prior Academic Performance, and United States Medical Licensing Examination Scores. Acad Med J Assoc Am Med Coll. 2019;94(3):364–70.

51. Cuddy MM, Swanson DB, Clauser BE. A multilevel analysis of examinee gender and USMLE step 1 performance. Acad Med J Assoc Am Med Coll. 2008 Oct;83(10 Suppl):S58-62.

52. Wakeford R, Denney M, Ludka-Stempien K, et al. Cross-comparison of MRCGP & MRCP(UK) in a database linkage study of 2,284 candidates taking both examinations: assessment of validity and differential performance by ethnicity. BMC Med Educ. 2015 Jan 16;15(1):1.

53. Hope D, Adamson K, McManus IC, et al. Using differential item functioning to evaluate potential bias in a high stakes postgraduate knowledge based assessment. BMC Med Educ. 2018 Apr 3;18(1):64.

54. Woolf K, Potts HWW, McManus IC. Ethnicity and academic performance in UK trained doctors and medical students: systematic review and meta-analysis. BMJ. 2011 Mar 8;342:d901.

55. Ross PT, Hart-Johnson T, Santen SA, et al. Considerations for using race and ethnicity as quantitative variables in medical education research. Perspect Med Educ. 2020 Oct;9(5):318–23.

56. Khine MS. Non-Cognitive Skills and Factors in Educational Success and Academic Achievement. In: Khine M.S., Areepattamannil S. (eds) Non-cognitive Skills and Factors in Educational Attainment. Contemporary Approaches to Research in learning Innovations. Rotterdam: SensePublishers; 2016. p. 3–9.

57. Gutman LM, Schoon I. The impact of non-cognitive skills on outcomes for young people. [Internet]. Institute of Education, London.; 2013 [cited 2020 May 19]. Available from: https://educationendowmentfoundation.org.uk/uploads/pdf/Noncognitive_skills_literature_review.pdf.

58. MacKenzie RK, Dowell J, Ayansina D, et al. Do personality traits assessed on medical school admission predict exit performance? A UK-wide longitudinal cohort study. Adv Health Sci Educ. 2017 May;22(2):365–85.

BMJ Open

Does performance at medical school predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination? A retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-046615.R1
Article Type:	Original research
Date Submitted by the Author:	23-May-2021

Complete List of Authors:	Ellis, Ricky; University of Aberdeen Institute of Applied Health Sciences; Nottingham University Hospitals NHS Trust, Urology Department Scrimgeour, Duncan; University of Aberdeen Institute of Applied Health Sciences; Aberdeen Royal Infirmary, Department of Colorectal Surgery Brennan, Peter ; Portsmouth Hospitals NHS Trust, Department of Oral and Maxillofacial Surgery Lee, Amanda; University of Aberdeen Institute of Applied Health Sciences, Department of Medical Statistics Cleland, Jennifer; Lee Kong Chian School of Medicine, Medical Education Research and Scholarship Unit (MERSU)

Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Health policy, Surgery
Keywords:	MEDICAL EDUCATION & TRAINING, SURGERY, Adult surgery < SURGERY

SCHOLARONE™ Manuscripts

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Does performance at medical school predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination? A retrospective cohort study

Ricky Ellis[1,2,3], Duncan SG Scrimgeour[1,3], Peter A Brennan[4], Amanda J Lee[5], Jennifer Cleland[6]

[1]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen and [2]Urology Department, Nottingham University Hospitals, Nottingham and [3]Department of Colorectal Surgery, Aberdeen Royal Infirmary, Aberdeen and [4]Department of Maxillo-Facial Surgery, Queen Alexandra Hospital, Portsmouth and [5]Department of Medical Statistics, Institute of Applied Health Sciences, University of Aberdeen, [6]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.

Correspondence to: Mr Ricky Ellis

Institute of Applied Health Sciences, University of Aberdeen,

Aberdeen, AB25 2ZD

Rickyellis@nhs.net

ABSTRACT

Background: Identifying predictors of success in post-graduate examinations can help guide the career choices of medical students and may aid early identification of trainees requiring extra support to progress in specialty training. We assessed whether performance on the Educational Performance Measurement (EPM) and Situational Judgement Test (SJT) used for selection into Foundation Training predicted success at the Membership of the Royal College of Surgeons (MRCS) examination.

Methods: This was a longitudinal, cohort study using data from the UK Medical Education Database (https://www.ukmed.ac.uk). UK medical graduates who had attempted Part A (*n*=2585) and Part B (*n*=755) of the MRCS between 2014-2017 were included. Chi-squared and independent t-tests were used to examine the relationship between medical school performance and socio-demographic factors with first-attempt success at MRCS Part A and B. Multivariate logistic regression was employed to identify independent predictors of MRCS performance.

Results: For every additional EPM decile point gained, the odds of passing MRCS increased by 55% for Part A (odds ratio 1.55 [95% confidence interval (CI) 1.48-1.61]) and 23% for Part B (1.23 [1.14-1.32]). For every point awarded for additional degrees in the EPM, candidates were 20% more likely to pass MRCS Part A (1.20 [1.13-1.29]) and 17% more likely to pass Part B (1.17 [1.04-1.33]). For every point awarded for publications in the EPM, candidates were 14% more likely to pass MRCS Part A (1.14 [1.01-1.28]). SJT score was not a statistically significant independent predictor of MRCS success.

Conclusion: This study has demonstrated the EPM's independent predictive power and found that medical school performance deciles are the most significant measure of predicting later success in the MRCS. These findings can be used by medical schools, training boards and workforce planners to inform evidence-based and contemporary selection and assessment strategies.

Key words: MRCS, UKFPO, situational judgement test, Medical School, Surgery, Postgraduate examinations

Strengths and limitations of this study

- This is the first study to investigate the relationship between medical school performance with performance at a high stakes UK postgraduate surgical examination.
- This is a large retrospective cohort study using the U.K. Medical Education database.

- This study examines whether performance on the Educational Performance Measurement (EPM) and Situational Judgement Test (SJT) used for selection into Foundation Training predicted success at the MRCS examination.
- Following previous studies, the relatively blunt measure of MRCS pass/fail results at first attempt was used as the primary outcome

INTRODUCTION

Progression through the United Kingdom (UK) medical education and training pathway is based on performance on a series of index assessments, starting with examination performance prior to entry to medical school and (typically) ending with respective Royal College Fellowship examinations. Each assessment is designed to ensure appropriate standards for stage of training and to ultimately safeguard patients.(1,2)

Performance at each stage also has implications on career progression. In the UK, doctors with higher academic scores during medical school are more likely to be offered their first choice of UK Foundation Programme (UKFP) training post on graduating.(3) Those with higher academic scores during medical school are also more likely to be offered a training place in a more competitive specialty.(4)

Studies have already demonstrated the validity of academic performance during medical school in predicting performance during Foundation Training,(5-7) although there is little research on the association between medical school performance and performance during specialty training in the UK. The seminal "Academic Backbone" paper by McManus et al. (2013) described how prior attainment is the best predictor of future performance in medical education.(8) However, that study was carried out before standardized markers of medical school performance were introduced (see later) and therefore may not represent contemporary patterns of performance. This is an important deficiency in the literature as research from other contexts indicates that examination results taken during and shortly after medical school predict later performance on board certification examinations, and patient complaints.(9-14) Furthermore, if early assessments do not predict later performance, then their fitness for purpose as markers of performance and their use as gateways for progression in training are questionable.

At the time of this study the UK did not have a national licensing examination for graduating doctors. Instead, performance during medical school is measured within schools by the Educational Performance Measure, or EPM.(15) The EPM is calculated out of 50 points and comprises three parts (Table 1); medical school performance decile (points are awarded depending on a student's final EPM decile; ranging from 34 points for the 10^{th} (lowest) decile to 43 points for students in the 1^{st} (highest) decile); additional degrees (0-5 points are awarded according to degree grade achieved); and publications (maximum 2 points; 1 point is awarded per publication). The EPM decile is calculated using multiple assessments of a student's knowledge and practical skills over time and by multiple assessments throughout medical school.(15) Points awarded for additional degrees and publications (0-7 in total) are described as Educational Achievements (EA). The selection process for the UKFP couples the EPM with a situational judgement test, or SJT (16–21) also scored out of 50 points. The UKFP SJT could be described as a type of 'procedural knowledge test'; assessing procedural knowledge about what to do in certain situations and how to do it.(21) The procedural knowledge being assessed by the UKFP SJT aligns with the behaviours and attitudes expected of doctors as described in the General Medical Council's (GMC) Good Medical Practice.(2) The graduate's combined EPM score plus SJT score out of 100 is their application score for the UKFP.(22)

Table 1. The components of the Educational Performance Measure used to quantify performance at medical school. The EPM is combined with a candidates SJT score to create their total UKFP selection score.

Educational Performance Measure (EPM)	**Points available**
Medical School Performance Decile	34-43
Additional Degrees	0-5
Publications	0-2
Maximum combined EPM score	**50**

Situational Judgment Test (SJT)	**Points available**
Test Score	50

United Kingdom Foundation Program (UKFP) Selection Score	**Points available**
EPM	50
SJT	50
Maximum UKFP Selection Score	**100**

The current study aimed to assess whether performance in medical school, EPM and SJT scores, could predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination. The MRCS examination is often taken by UK trainees during Foundation and Core surgical training years and comprises of two parts: Part A, a written examination with two papers and Part B, an objective structured clinical examination (OSCE).(23,24) The MRCS is a high-stakes postgraduate assessment that is used as a gateway for applications for higher surgical training and is itself a good predictor of future surgical training outcomes.(25-27)

Given performance at medical school and success in postgraduate assessments is related to socio-demographic factors as well as academic ability, regression models were adjusted for socio-demographic factors known to be associated with MRCS success.(3,23,25,28–30) These included gender, ethnicity and graduate status on entry to medical school. This analysis is timely given policy drivers in the UK to ensure that medical school and postgraduate assessments are fair,(1) the pending imposition of a once-off high-stakes test, the Medical Licensing Assessment (MLA),(31) and proposals to exclude Educational Achievements from the EPM score used in UKFP selection from 2023.(32)

Use of linked individual-level data from the UK Medical Education Database (UKMED: https://www.ukmed.ac.uk/) enabled a national-level analysis, drawing on data from sources including medical school assessment, Foundation Programme selection and postgraduate assessment outcomes.(33)

METHODS

A longitudinal retrospective cohort study was conducted on UK medical graduates who had attempted either the Part A (written) or the Part B (clinical) MRCS examination from April 2014 to May 2017.

The UK Medical Education Database (UKMED: https://www.ukmed.ac.uk/) was used to access linked data from

UK medical schools and the four Royal Colleges of Surgeons in the UK and Ireland. All counts have been rounded according to Higher Education Statistics Agency (HESA) data standards to ensure person-level anonymity.(34)

The following data were extracted: self-declared gender, self-reported ethnicity demographics and graduation status at the time of entry to medical school, medical school Educational Performance Measure decile, additional degree and EPM publication scores, SJT score and MRCS Part A and B first attempt result. Figure 1 shows the flow of data through the study. Candidate first attempt results were used as they have been shown to be the best predictor of future performance in postgraduate examinations.(35)

Figure 1. Data flow through study.

Except for SJT and EPM scores, all variables were subsequently dichotomized. Graduation status was defined as “yes” if candidates had obtained a degree prior to entering medicine. Self-declared ethnicity was coded as “white” or “non-white” as used in similar studies to enable powered analysis of smaller cohorts.(25,26) Part A and B MRCS performance was categorized as “pass”’or “fail” at first attempt.

Statistical analysis

All analyses were conducted using SPSS® v22.0 (IBM, Armonk, New York, USA). A chi-squared test was initially employed to determine any associations with first attempt MRCS pass/fail outcomes. The relationship between SJT, EPM decile, additional degrees, EPM publication scores and Part A and Part B MRCS first attempt success was examined using independent t-tests since the distribution of scores was normal. Correlation coefficients were calculated for Foundation Programme selection scores and the MRCS Part A (the written component of the MRCS examination) score relative to pass mark.

Logistic regression models were developed to identify predictors of success at MRCS at first attempt that were independent of other performance measures used in UKFP selection. Further regression models were developed to identify predictors of MRCS success, that were independent of other performance metrics and socio-demographic factors known to be associated with MRCS performance. While doctors are not selected for the UKFP based on socio-demographic factors, adjusting for these known predictors of MRCS success ensured that regression models were adjusted for these potential confounding factors and are therefore more applicable in real life. Potential interactions between significant predictors were also examined.

The Intercollegiate Committee for Basic Surgical Examinations (ICBSE) and its Internal Quality Assurance Subcommittee, which monitors standards and quality, approved this study. The highest standards of security, governance and confidentiality were ensured when storing, handling and analysing identifiable data.

Patient and Public involvement

No patients or public were involved in this study.

RESULTS

Part A MRCS

A total of 3000 UK medical graduates attempted Part A MRCS between April 2014 and May 2017. Of these 2585 had matched EPM and SJT data. 50% (n=1280) passed Part A MRCS at their first attempt. 63% of candidates (n=1635), were male, 56% were white (n=1435), and 81.5% had not undertaken a prior degree before entering medicine (n=2105). Mean (standard deviation [SD]) total EPM and SJT scores for candidates who had attempted Part A MRCS were 41.6 (3.86) and 39.4 (3.54) respectively.

Pass rates for Part A MRCS by gender, ethnicity and graduate on entry to medicine status are shown in Table 2. Differences in pass rates were statistically significant for: gender (54.9% males *vs.* 40.5% females, $P<0.001$), ethnicity (54.2% white *vs.* 44.1% non-white, $P<0.001$) and graduate status (50.7% no prior degree *vs.* 44.7% prior degree, $P=0.017$).

Table 2. Univariate analysis of MRCS first attempt pass rates by gender, ethnicity and graduation status for UK medical graduates.

Variable	Part A (*n*= 2585)		Part B (*n*= 755)	
	Pass *n*=1280	**Fail** *n*=1305	**Pass** *n*=575	**Fail** *n*=180
Gender				
Male	54.9% (895/1635)	45.1% (740/1635)	75.9% (380/500)	24.1% (120/500)
Female	40.5% (385/950)	59.5% (565/950)	77.1% (195/255)	22.9% (60/255)
p-value	< 0.001		0.719	
Ethnicity				
White	54.2% (780/1435)	45.8% (660/1435)	82.0% (350/430)	18.0% (75/430)
Non-white	44.1% (500/1135)	55.9% (635/1135)	68.8% (225/325)	31.2% (105/325)
Missing	*n*=10		*n*=0	
p-value	< 0.001		< 0.001	
Graduate on entry to medicine				
No	50.7% (1070/2105)	49.3% (1040/2105)	77.8% (495/635)	22.2% (140/635)
Yes	44.7% (210/480)	55.3% (265/480)	68.3% (80/120)	31.7% (40/120)
p-value	0.017		0.025	

Note. All p-values presented are from chi-squared analysis. MRCS, Membership of the Royal College of Surgeons.

Univariate analysis of EPM and SJT scores are shown in Table 3. Candidates who passed Part A MRCS at first attempt had performed better in their SJT (Mean 40.0 [SD 3.3] *vs.* 38.9 [3.7], *P*<0.001) and had scored higher for their total EPM

(43.6 [3.3] *vs.* 39.8 [3.4], *P*<0.001) compared to those who failed at first attempt. Figure 2 shows the relative increase in mean MRCS Part A pass rates at first attempt according to candidate EPM decile.

Table 3. Univariate analysis of EPM scores, SJT scores and MRCS Part A and Part B first attempt success.

Variable		MRCS Part A			MRCS Part B		
		Pass *n*=1280	**Fail** *n*=1305	**P-value**	**Pass** *n*=575	**Fail** *n*=180	**P-value**
EPM Decile	Mean S.D.	40.1 2.4	37.3 2.5	<0.001	39.7 2.6	38.1 2.5	<0.001
EPM Degree score	Mean S.D.	2.8 1.5	2.1 1.6	<0.001	2.7 1.5	2.3 1.6	0.001
EPM Publications	Mean S.D.	0.7 0.9	0.5 0.8	<0.001	0.9 0.9	0.7 0.8	0.042
SJT	Mean S.D.	40.0 3.3	38.9 3.7	<0.001	40.2 3.2	39.4 4.2	0.010

Note. MRCS, Membership of the Royal College of Surgeons. EPM, Educational Performance Measure. SJT, Situational Judgment Test, N, number of candidates. S.D., Standard Deviation. *P*, *P* Value.

Figure 2. Relative increase in mean MRCS pass rates at first attempt according to candidate Educational Performance Measure (EPM) decile (1st EPM decile indicates the highest achieving candidates and 10th decile, the lowest achieving candidates).

Table 4 shows correlation coefficients between each Foundation Programme selection score and MRCS Part A. According to Cohen's guidelines (36) EPM degree score, EPM publication score and the SJT show statistically significant weak positive correlation with Part A scores. Total EPM and EPM decile show statistically significant strong correlations with MRCS Part A.

Table 4. Correlation coefficients between Foundation Programme selection scores and MRCS Part A Scores (*n*= 2585).

	Correlation Coefficient	P-value
Total EPM	0.57	<0.001
EPM Decile*	0.59	<0.001
EPM Degree Score	0.27	<0.001
EPM Publications	0.17	<0.001
SJT	0.23	<0.001

*Spearman's Rho coefficient.

Table 5 shows the odds ratios (OR) and 95% confidence intervals (CI) for independent predictors of passing Part A MRCS at first attempt. Odds ratios were similar for UKFP selection metrics when multivariate analysis included socio-demographic predictors of MRCS success. EPM decile, EPM degree and EPM publication scores were predictors of MRCS success independent of other selection metrics and socio-demographic factors. For every additional EPM decile, the odds of passing Part A MRCS at first attempt increased by 55% (OR 1.55, 95% CI 1.48 to 1.61), for every additional EPM degree point awarded, the odds of passing Part A on first attempt increased by 20% (OR 1.20, 95% CI 1.13 to 1.29) and for every additional EPM publication point awarded, the odds of passing Part A on first attempt increased by 14% (OR 1.14, 95% CI 1.01 to 1.28). SJT score was not found to independently predict Part A first attempt success (*P*=0.177). There was a statistically significant interaction between ethnicity and gender in the final Part A MRCS regression model with white males more likely to pass, *P*=0.002. MRCS candidates who entered medical school without a prior degree were more than twice as likely to pass Part A compared to those who entered medical school as graduates (OR 2.23, 95% CI 1.73 to 2.87).

Table 5. Predictors of pass at first attempt at Part A and Part B MRCS for UK medical graduates on multivariate analysis. Odds ratios are unadjusted and then adjusted for socio-demographic predictors of MRCS success.

	Part A		**Part B**	
Variable	**Unadjusted Odds ratio (95% CI)** *n*= 2585	**Adjusted Odds ratio (95% CI)** *n*=2570	**Unadjusted Odds Ratio (95% CI)** *n*=755	**Adjusted Odds Ratio (95% CI)** *n*= 755
Educational Performance Measure Decile	1.50 (1.45 to 1.56)	1.55 (1.48 to 1.61)	1.23 (1.15 to 1.32)	1.23 (1.14 to 1.32)
Educational Performance Measure Additional Degree points	1.16 (1.09 to 1.23)	1.20 (1.13 to 1.29)	1.09 (0.97 to 1.23)	1.17 (1.04 to 1.33)
Educational Performance Measure Publication Score	1.17 (1.04 to 1.31)	1.14 (1.01 to 1.28)	1.04 (0.85 to 1.28)	1.02 (0.83 to 1.26)
Situational Judgement Test	1.02 (0.99 to 1.04)	1.02 (0.99 to 1.05)	1.04 (0.99 to 1.09)	1.02 (0.97 to 1.10)
Gender* Males vs. Females	-	1.60 (1.17 to 2.10)	-	0.97 (0.66 to 1.43)
Ethnicity* White vs. Non-White	-	0.70 (0.51 to 0.96)	-	1.86 (1.29 to 2.69)

Ethnicity*Gender	-	1.97 (1.33 to 2.91)	-	-
Graduate on entry into medicine Non-Graduates vs. Graduates	-	2.23 (1.73 to 2.87)	-	2.54 (1.57 to 4.13)

MRCS, Membership of the Royal College of Surgeons; CI, Confidence Interval. *P=0.002 for interaction between Ethnicity and Gender in MRCS Part A Model.

Part B MRCS

In total, 755 of the Part A study cohort (n=2585) attempted MRCS Part B at a later date. 76.3% (n=575) of candidates passed Part B MRCS at first attempt. Unsurprisingly the demographics for Part B MRCS were similar to those observed for Part A MRCS candidates; 67% of candidates were male (n=500), 57% were white (n=430), and 84% had not undertaken a previous degree (n=635). The mean (SD) total EPM and SJT scores for UK graduates who had attempted Part B MRCS were 42.8 (3.67) and 40.0 (3.47) respectively.

Pass rates for Part B MRCS by gender, ethnicity and graduate on entry to medicine status are shown in Table 2. There was no significant difference in Part B MRCS first attempt pass rates between males and females (75.9% *vs.* 77.1% respectively P=0.719). Differences in pass rates were statistically significant for ethnicity (82.0% white *vs.* 68.8% non-white, *P*<0.001) and graduate status (77.8% no prior degree *vs.* 68.3% prior degree, *P*=0.025).

Univariate analysis of EPM and SJT scores are shown in Table 3. Those who passed Part B MRCS at first attempt had performed better in their SJT compared to those who failed at first attempt (40.2 [3.2] *vs.* 39.4 [4.2], *P*<0.010). Similarly, candidates who passed Part B at first attempt had scored higher in their total EPM (43.3 [3.6] *vs.* 41.1 [3.4], *P*<0.001). Figure 2 shows MRCS Part B performance according to EPM decile score. The overall trend reveals a relative increase in mean MRCS Part B pass rates at first attempt according to candidate EPM decile.

Table 5 shows the logistic regression models for independent predictors of Part B MRCS first attempt. EPM decile and EPM degree scores were statistically significant predictors of MRCS success independent of other selection metrics and socio-demographic factors. For every additional EPM decile the odds of passing Part B MRCS at first attempt increased by 23% (OR 1.23, 95% CI 1.14 to 1.32) and for every additional EPM degree point awarded, the odds of passing Part B MRCS at first attempt increased by 17% (OR 1.17, 95% CI 1.04 to 1.33). Neither SJT score or EPM publication scores were found to be independent predictors of Part B success at first attempt after adjusting for UKFP selection metrics and socio-demographic factors (*P*=0.429 and *P*=0.849 respectively).

White UK medical graduates were nearly twice as likely to pass Part B MRCS at first attempt compared to non-white candidates (OR 1.86, 95% CI 1.29 to 2.69). Candidates who had not undertaken a previous degree before entering medicine were more than twice as likely to pass Part B MRCS compared to those who had undertaken a prior degree (OR 2.54, 95% CI 1.57 to 4.13). There were no statistically significant interactions between any of the Part B MRCS variables in the adjusted regression model.

DISCUSSION

Educational Performance Measure - Deciles

We assessed the predictive validity of UKFP selection measures, the SJT and EPM, against the MRCS examination which is known to be a good predictor of future surgical training outcomes.(25-27) We found that EPM deciles predicted success at both Part A (written) and Part B (OSCE) of the MRCS independent of other UKFP selection scores and socio-demographic factors. For every incremental EPM decile, candidates were significantly more likely to pass both MRCS Part A and Part B. Reassuringly, the predictive value of EPM deciles were not significantly altered when adjusting for gender, ethnicity and graduate status, indicating that very little of the association that exists between FP selection scores and MRCS performance is explained by these socio-demographic factors. Our results add to previous studies which found that the EPM predicts performance during Foundation training,(6,7) and provides assurance that UK medical school assessments appropriately gauge student competence and readiness for practice.

A key limitation of the EPM decile score as a selection tool is its ranking of medical school graduates at a local rather than national level. Each medical school ranks their cohort of graduates internally into 10 equal groups (deciles) based on performance in a number of assessments taken over the duration of the medical course. Students are therefore ranked against their peers within each medical school, potentially penalising high-achieving individuals that study at more competitive schools, resulting in a lower decile score than if those individuals studied at schools with a less competitive cohort. Given that assessment also varies significantly in 'volume, type and intensity' between medical schools, concerns have been raised that students of equal proficiency may fall into different EPM deciles across schools due to differences in assessment rather than ability. (37) Furthermore, the number of assessments used and scoring for each assessment varies considerably between schools which can limit the range of scores used for decile ranking, reducing the spread of candidates.(38)

Concerns regarding the impact of variation in local assessment and ranking have resulted in demand for the MLA in the UK. A national MLA is argued to provide a potentially more robust method of ranking medical graduates nationally and may also contribute to standard setting for education across medical schools. The MLA's impending introduction has been met with a mixed response with some arguing that a single high-stakes exit examination is not as valuable as multiple local assessments over a number of years and may also result in schools teaching to pass instead of teaching to practice medicine.(37,39) However, a one-off high-stakes examination on completion of medical school reflects the use of assessments throughout postgraduate medical training. The predictive and incremental validity of the new MLA must be scrutinised to justify its financial cost and its burden on both students and the medical education system.

Despite the limitations and potential shortcomings of the EPM decile scoring system that is currently being used, it appears to achieve its intended function for UKFP selection. It differentiates between candidates by ability and demonstrates the ability to predict postgraduate performance.(6,40) These data support its predictive validity and ongoing use as a UKFP selection tool. How the EPM, a local-level assessment, will sit alongside the proposed MLA, remains to be seen.

Our findings also align with the "academic backbone" concept proposed by McManus et al. (2013); an idea that in medical education, current learning and achievement is dependent upon attainment at earlier stages.(8) This can be summarised simply as: medical students who are high achievers remain high achievers. Candidates ranked in the top deciles perform better at MRCS. Those who perform best in MRCS are more likely to achieve a specialty training (ST3) post at national selection, are more likely to progress through training with satisfactory Annual Review of Competence Progression (ARCP) outcomes and are more likely to succeed at first attempt in their surgical fellowship (FRCS) examinations.(25–27) The nature of the outcome measures have changed since the seminal study of McManus et al. in 2013,(8) but the principles have not: the road to success for those who wish to pursue a successful career in surgery begins early.

Educational Performance Measure - Educational Achievements (EAs)

Points awarded in the EPM for additional degrees predict success at MRCS independent of other UKFP selection measures and socio-demographic factors. While points awarded for additional publications independently predict success in MRCS Part A, they were not an independent predictor of success in MRCS Part B. Correlations between MRCS Part A scores and EA points were considerably weaker than the correlation with EPM decile scores. It also appears that EPM decile scores are largely responsible for the strength of the correlation seen between EPM total and MRCS Part A scores. These results are timely and relevant given the recent announcement that points awarded for EA will be excluded from the EPM scoring system for UKFP selection from 2023.(32)

Points awarded for EA in the EPM undoubtedly play a role in increasing the spread of applicant scores when combined with EPM decile and SJT points for UKFP selection.(40) However, there is evidence of increasing EA point inflation over recent years with the number of applicants earning EA points increasing from 30% to 70%.(41) Given this EA point inflation, it is possible that correlations between EA scores and MRCS performance found in our data may be higher than those seen in cohorts that have graduated from medical school more recently. It is clear that over time the ability of EA points to differentiate candidates will diminish, but the financial barriers to success in medicine that these may cause would persist, with students from more affluent backgrounds being in a position to 'pay for points' by studying an intercalated degree. Indeed, given the recent drive to widen access to medicine it would appear contradictory for selection tools to encourage students to take on the significant financial burden of an intercalated degree that is not necessary for the practice of medicine, and does not necessarily improve patient care. Studying an intercalated degree does undoubtedly have many advantages that would 'enrich the student experience'(42), but students should not be penalised in their national ranking if uninterested or unable to afford to do so.

Overall, it could be argued that the limited predictive value of EA points found in this study and others (6) does not outweigh their potential to limit the score and subsequent ranking of applicants' that are unable to afford to undertake an intercalated degree.

Foundation Programme Situational Judgement Test

Candidates who passed both parts A and B of the MRCS at first attempt scored higher in their SJT than candidates who failed, and there was a statistically significant positive correlation with Part A scores. However, the SJT did not independently predict MRCS success after adjusting for EPM scores and socio-demographic factors, displaying no significant incremental value over and above the predictive value of EPM decile scores. It is important to consider this finding in relation to the premise behind SJTs.

The Foundation Programme (FP) SJT is based on a job analysis of being a Foundation Doctor.(7) Significant correlation between SJT and EPM scores between schools has been identified.(6,37,43) Additionally both SJT and EPM scores are independently associated with the odds of successful completion of the FP, and SJT score offers a degree of incremental predictive validity over that provided by the EPM deciles, suggesting that it is capturing additional, relevant, information on applicants, as intended.(6,40) Research suggests that the FP SJT does what it was designed to do as well as succeeding in increasing the spread of candidates being ranked for foundation training posts (the arguments as to how it should be weighed in the FP selection process are outside the scope of this paper but we direct readers with an interest to other studies).(6,44,45) It was not designed to select for specialty training: where specialty training programmes use SJTs for selection, these have been designed specifically against the role of a trainee/resident in that speciality.(46-48) Given this, in retrospect it is unsurprising that the FP SJT does not independently predict performance on a post-graduate examination that tests the clinical knowledge, skills and professional attitudes expected of surgical trainees.

Strengths and Weaknesses

The current study is one of the first to use the UKMED to examine the associations between medical school performance and FP SJT outcomes on success at a high-stakes postgraduate surgical examination. The UKMED enabled a nationwide, multi-cohort analysis and our breakdown of the FP selection process into EPM scores and SJT allowed us to look separately at academic attainment and other factors.

There are some limitations of the study. Firstly, although candidates can take Parts A and B MRCS on multiple occasions, we used candidate first attempt results as the best predictor of future performance.(35) We often used the relatively blunt outcome measure of pass/fail as this is what is meaningful to those sitting the MRCS, and has been used in previous studies looking at factors that predict performance in the MRCS.(25)

Self-declared ethnicity data were combined into two discrete categories to maximise power when analysing smaller cohorts, rather than this being an ethical or social decision. Regression analyses were adjusted for known socio-demographic predictors of MRCS success, but these were not the main focus of the current paper. We are currently undertaking further analyses to characterise group-level attainment differences that have been identified.(49) Finally, the current analysis was based on retrospective quantitative data. A prospective study would have allowed us to examine more variables related to progression and attainment in surgical training. For example, being good at passing exams is linked to academic ability, but the wider education literature makes clear that non-cognitive factors such as motivation, time management and resilience are also relevant to performance.(50,51) If appropriate measures could be identified,(52) it would be interesting to compare graduates and MRCS candidates on these factors.

CONCLUSION

Success at first attempt of MRCS Part A and B can be predicted from medical school performance (EPM decile score) but not from the Foundation Programme SJT score. Put simply, medical students who do well in terms of medical school examination performance remain strong performers later on in their careers. These results may help to guide career choices for students and can be used by training institutions to inform evidence-based and contemporary selection and assessment strategies.

ABBREVIATIONS

ARCP: Annual Review of Competence Progression

EA: Educational Achievements

EPM: Educational Performance Measure

FRCS: Fellowship of the Royal College of Surgeons Examinations

GMC: General Medical Council

HESA: Higher Education Statistics Agency

MLA: United Kingdom Medical Licensing Examination

MRCS: Intercollegiate Membership of the Royal College of Surgeons Examinations

SJT: Situational Judgement Test

UKFP: United Kingdom Foundation Program

USMLE: United States Medical Licensing Examinations

ACKNOWLEDGEMENTS

The authors would like to acknowledge Iain Targett at the Royal College of Surgeons of England, for his help with data collection and Gregory Ayre from the Intercollegiate Committee for Basic Surgical Examinations for their support during this project. Our thanks to members of the UKMED Research Group who provided useful feedback on an earlier version of this manuscript, and whose comments helped refine the paper. The authors would also like to acknowledge Daniel Smith for his help with the UKMED database. Data Source: UK Medical Education Database ("UKMED"). UKMEDP043 extract generated on 25/07/2018. We are grateful to UKMED for the use of these data. However, UKMED bears no responsibility for their analysis or interpretation. The data includes information derived from that collected by the Higher Education Statistics Agency Limited ("HESA") and provided to the GMC ("HESA Data"). Source: HESA Student Records 2007/2008 to 2015/2016. Copyright Higher Education Statistics Agency Limited. The Higher Education Statistics Agency Limited makes no warranty as to the accuracy of the HESA Data, cannot accept responsibility for any inferences or conclusions derived by third parties from data or other Information supplied by it.

FUNDING

Royal College of Surgeons of England, Royal College of Surgeons of Edinburgh, Royal College of Surgeons of Ireland and Royal College of Physicians and Surgeons of Glasgow (Award/Grant number is not applicable).

AVAILABILITY OF DATA AND MATERIALS

The dataset used in this study was acquired from the UK Medical Education Database and is held in Safe Haven. Data access requests must be made to UKMED. Full information for applications can be found at https://www.ukmed.ac.uk.

CONFLICTS OF INTEREST

None to declare.

ETHICAL APPROVAL

No formal ethical approval was required for this study of existing UKMED data. UKMED has received ethics exemption for projects using exclusively UKMED data from Queen Marys University of London Ethics of Research Committee on behalf of all UK medical schools (https://www.ukmed.ac.uk/documents/UKMED_research_projects_ethics_exemption.pdf). The Intercollegiate Committee for Basic Surgical Examinations (ICBSE) and its Internal Quality Assurance Subcommittee, which monitors MRCS standards, research and quality, approved this study.

CONTRIBUTORSHIP

RE and DS wrote the first draft of the manuscript. RE and DS performed statistical analyses with AL's supervision. RE, DS, PB,AL and JC reviewed and edited the manuscript. JC led the study proposal for access to UKMED data. All authors approved final draft of the manuscript.

REFERENCES

1. Promoting excellence: standards for medical education and training [Internet]. General Medical Council; 2015 [cited 2021 May 21]. Available from: www.gmc-uk.org/education/standards.asp

2. Good Medical Practice [Internet]. General Medical Council; 2013 [cited 2021 May 21]. Available from: www.gmc-uk.org/guidance.

3. Kumwenda B, Cleland JA, Prescott GJ, et al. Relationship between sociodemographic factors and selection into UK postgraduate medical training programmes: a national cohort study. BMJ Open. 2018 Jun;8(6):e021329.

4. Kumwenda B, Cleland J, Prescott G, et al. Relationship between sociodemographic factors and specialty destination of UK trainee doctors: a national cohort study. BMJ Open. 2019 Mar;9(3):e026961.

5. Patterson F, Ashworth V, Murray H, et al. Analysis of the Situational Judgement Test for Selection to the Foundation Programme 2013. Technical Report. [Internet]. Work Psychology Group; 2013 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2017/04/fy1-sjt-2013-technical-report.pdf

6. Smith DT, Tiffin PA. Evaluating the validity of the selection measures used for the UK's foundation medical training programme: a national cohort study. BMJ Open. 2018 Jul 1;8(7):e021918.

7. Medical Schools Council on behalf of the cross-stakeholder Project Group. Improving Selection to the Foundation Programme Final Report of the Parallel Recruitment Exercise. [Internet]. 2011 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2017/04/isfp-final-report-appendices.pdf

8. McManus I, Woolf K, Dacre J, et al. 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 Dec;11(1):242.

9. de Virgilio C, Yaghoubian A, Kaji A, et al. Predicting performance on the American Board of Surgery qualifying and certifying examinations: a multi-institutional study. Arch Surg Chic Ill 1960. 2010 Sep;145(9):852–6.

10. Tamblyn R, Abrahamowicz M, Brailovsky C, et al. Association between licensing examination scores and resource use and quality of care in primary care practice. JAMA. 1998 Sep 16;280(11):989–96.

11. Tamblyn R, Abrahamowicz M, Dauphinee D, et al. Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. JAMA. 2007 Sep 5;298(9):993–1001.

12. McCaskill QE, Kirk JJ, Barata DM, et al. USMLE step 1 scores as a significant predictor of future board passage in pediatrics. Ambul Pediatr Off J Ambul Pediatr Assoc. 2007 Apr;7(2):192–5.

13. Wenghofer E, Klass D, Abrahamowicz M, et al. Doctor scores on national qualifying examinations predict quality of care in future practice. Med Educ. 2009 Dec;43(12):1166–73.

14. Norcini JJ, Boulet JR, Opalek A, et al. The relationship between licensing examination performance and the outcomes of care by international medical school graduates. Acad Med J Assoc Am Med Coll. 2014 Aug;89(8):1157–62.

15. Educational Performance Measure (EPM) 2020 Framework [Internet]. UK Foundation Programme; 2020 [cited 2021 May 21]. Available from: https://foundationprogramme.nhs.uk/resources/

16. Patterson F, Driver R. Situational Judgement Tests (SJTs). In: Patterson F, Zibarras L, editors. Selection and Recruitment in the Healthcare Professions [Internet]. Springer International Publishing; 2018 [cited 2021 May 21]. p. 79–112. Available from: http://link.springer.com/10.1007/978-3-319-94971-0_4

17. Patterson F, Knight A, Dowell J, et al. How effective are selection methods in medical education? A systematic review. Med Educ. 2016 Jan;50(1):36–60.

18. Mcdaniel M, Morgeson F, Finnegan E, et al. Use of Situational Judgment Tests to Predict Job Performance. J Appl Psychol. 2001 Sep 1;86:730–40.

19. McDaniel MA, Hartman NS, Whetzel DL, et al. Situational Judgment Tests, Response instructions and Validity: A Meta-Analysis. Pers Psychol. 2007 Mar;60(1):63–91.

20. Lievens F, Peeters H, Schollaert E. Situational judgment tests: a review of recent research. Pers Rev. 2008 Jun 6;37(4):426–41.

21. Webster ES, Paton LW, Crampton PES, Tiffin PA. Situational judgement test validity for selection: A systematic review and meta-analysis. Med Educ. 2020 Oct;54(10):888–902.

22. UK Foundation Programme 2020 Applicants Handbook. [Internet]. 2020 [cited 2021 May 21]. Available from: https://www.foundationprogramme.nhs.uk/wp-content/uploads/sites/2/2019/10/UKFP-2020-Applicants-Handbook-.pdf

23. Scrimgeour DSG, Cleland J, Lee AJ, et al. Which factors predict success in the mandatory UK postgraduate surgical exam: The Intercollegiate Membership of the Royal College of Surgeons (MRCS)? Surg J R Coll Surg Edinb Irel. 2018 Aug;16(4):220–6.

24. Scrimgeour D, Cleland J, Lee A, Brennan PA. When is the best time to sit the MRCS examination? BMJ. 2017 Feb 6;356:j461.

25. Scrimgeour DSG, Cleland J, Lee AJ, et al. Prediction of success at UK Specialty Board Examinations using the mandatory postgraduate UK surgical examination. BJS Open. 2019 Dec;3(6):865–71.

26. Scrimgeour D, Brennan P, Griffiths G, et al. Does the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination predict 'on-the-job' performance

during UK higher specialty surgical training? Ann R Coll Surg Engl. 2018 Nov;100(8):669–75.

27. Scrimgeour DSG, Cleland J, Lee AJ, et al. Impact of performance in a mandatory postgraduate surgical examination on selection into specialty training: BJS Open. 2017 Oct;1(3):67–74.

28. Richens D, Graham TR, James J, et al. Racial and Gender Influences on Pass Rates for the UK and Ireland Specialty Board Examinations. J Surg Educ. 2016 Jan 1;73(1):143–50.

29. Puddey IB, Mercer A, Carr SE. Relative progress and academic performance of graduate vs undergraduate entrants to an Australian medical school. BMC Med Educ. 2019 May 22;19(1):159.

30. Woolf K, Cave J, Greenhalgh T, et al. Ethnic stereotypes and the underachievement of UK medical students from ethnic minorities: qualitative study. BMJ. 2008;337(a1220).

31. Medical Licensing Assessment [Internet]. General Medical Council; 2020 [cited 2021 May 21]. Available from: https://www.gmc-uk.org/education/medical-licensing-assessment

32. For UKFP2023 Entry – Update on Educational Achievements [Internet]. UK Foundation Programme; 2020 [cited 2021 May 21]. Available from: https://healtheducationengland.sharepoint.com/:w:/s/UKFPOT/EYd3TEAU749MpYfS4tWDEO8Bx-kdOsDmPs73uTDAH-xzxA?e=d3HdRi

33. Dowell J, Cleland J, Fitzpatrick S, et al. The UK medical education database (UKMED) what is it? Why and how might you use it? BMC Med Educ. 2018 Dec;18(1):6.

34. Rounding and suppression to anonymise statistics | HESA [Internet]. [cited 2021 May 21]. Available from: https://www.hesa.ac.uk/about/regulation/data-protection/rounding-and-suppression-anonymise-statistics

35. McManus I, Ludka K. Resitting a high-stakes postgraduate medical examination on multiple occasions: nonlinear multilevel modelling of performance in the MRCP(UK) examinations. BMC Med. 2012 Jun 14;10(1):60.

36. Hemphill JF. Interpreting the magnitudes of correlation coefficients. Am Psychol. 2003;58(1):78–9.

37. Devine OP, Harborne AC, McManus IC. Assessment at UK medical schools varies substantially in volume, type and intensity and correlates with postgraduate attainment. BMC Med Educ. 2015 Sep 11;15(1):146.

38. Improving Selection to the Foundation Programme - Final Report [Internet]. Medical Schools Council on behalf of the cross-stakeholder Project Group; 2011 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2017/04/isfp-final-report-appendices.pdf

39. Noble ISG. Are national qualifying examinations a fair way to rank medical students? No. BMJ. 2008 Aug 22;337:a1279.

40. Patterson F, Kerrin M, Edwards H, et al. Validation of the F1 Selection Tools [Internet]. Work Psychology Group on behalf of Health Education England; 2015 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2018/06/validation_of_the_f1_selection_tools_report_final_for_publication-1.pdf

41. For UKFP2023 Entry – Update on Educational Achievements (December 2020 Follow-Up) [Internet]. UK Foundation Programme; 2020 [cited 2021 May 21]. Available from: https://healtheducationengland.sharepoint.com/:b:/g/UKFPO/EYiV65BOKgpJsu4qLnakqPIBNdNM9HQ7vl4GMMALeCP4Dg?e=ZTVjgY

42. MSC Statement on the UKFPO decision to remove Educational Achievements from the Foundation ranking process [Internet]. Medical Schools Council [cited 2021 May 21]. Available from: https://www.medschools.ac.uk/news/msc-statement-on-the-ukfpo-decision-to-remove-educational-achievements-from-the-foundation-ranking-process

43. McManus IC, Harborne AC, Horsfall HL, et al. Exploring UK medical school differences: the MedDifs study of selection, teaching, student and F1 perceptions, postgraduate outcomes and fitness to practise. BMC Med. 2020 May 14;18(1):136.

44. Najim M, Rabee R, Sherwani Y, et al. The situational judgement test: a student's worst nightmare. Adv Med Educ Pract. 2015 Oct 5;6:577–8.

45. Mclachlan JC, Illing J. An Independant review of events relating to the use of Situational Judgement Tests in the selection for Foundation process 2013. Commissioned by Health Education England [Internet]. 2013 [cited 2021 May 21]. Available from: https://www.medschools.ac.uk/media/1932/hee-independent-review-of-the-sjt.pdf

46. Patterson F, Rowett E, Hale R, et al. The predictive validity of a situational judgement test and multiple-mini interview for entry into postgraduate training in Australia. BMC Med Educ. 2016 Dec;16(1):87.

47. Applicant Guidance. General Practice: Multi-Specialty Recruitment Assessment (MSRA) [Internet]. Health Education England; [cited 2020 May 19]. Available from: https://gprecruitment.hee.nhs.uk/Resource-Bank/Recruitment-Documents-Forms

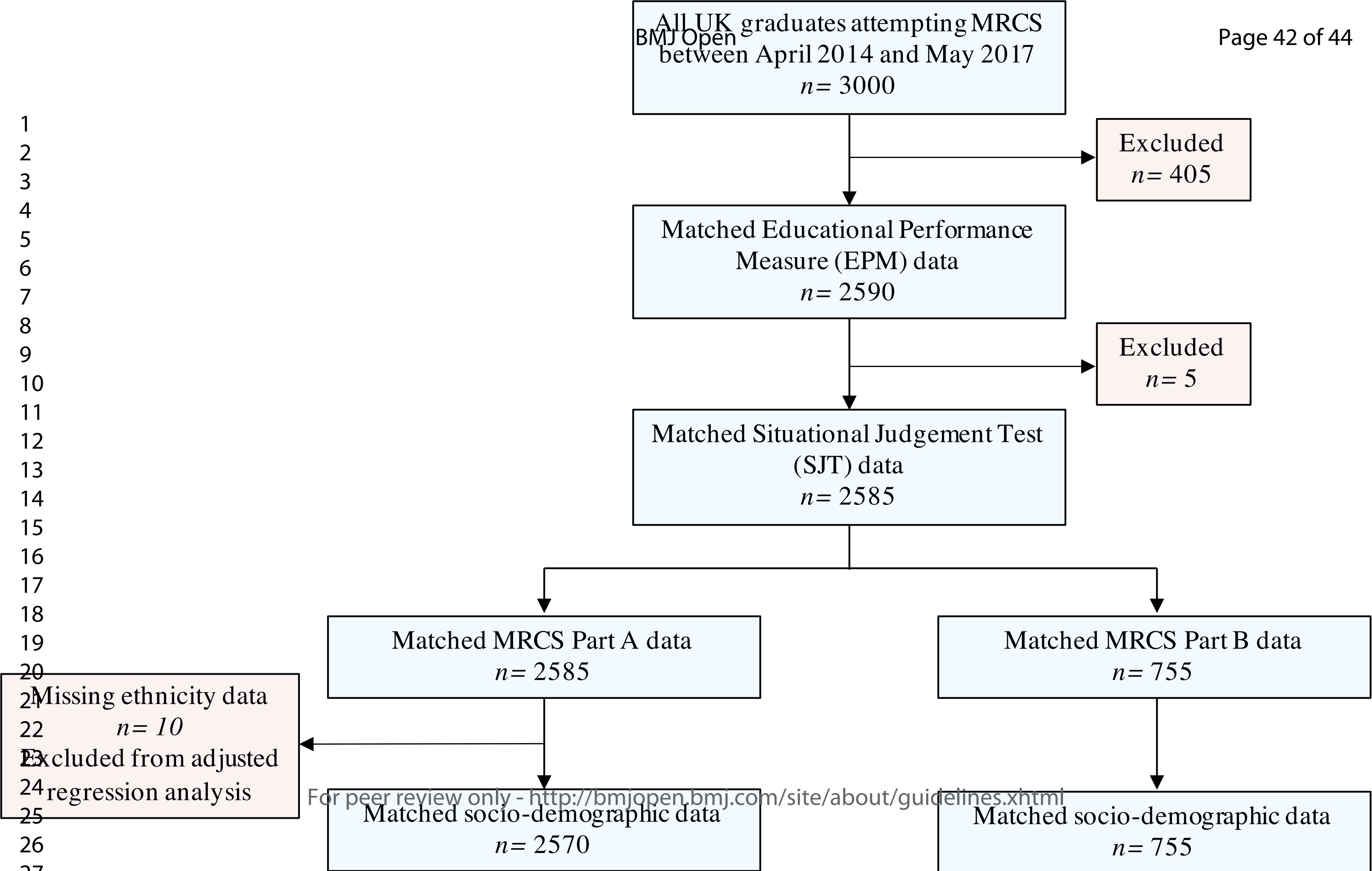
48. Multi-Specialty Recruitment Assessment (MSRA) – Test Blueprint & Information [Internet]. Work Psychology Group on behalf of Health Education England; 2019 [cited 2021 May 21]. Available from: https://gprecruitment.hee.nhs.uk/Resource-Bank/Recruitment-Documents-Forms

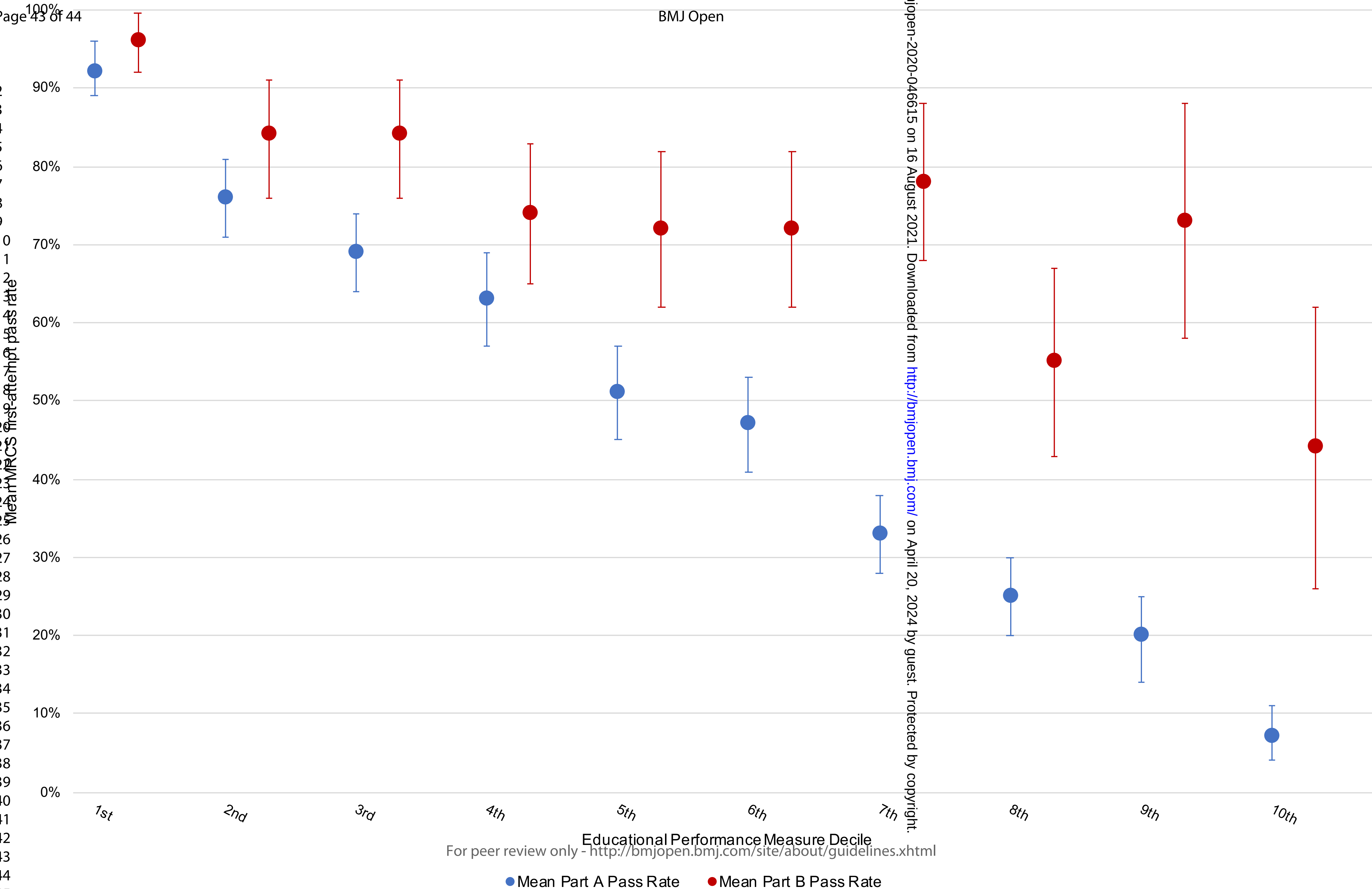
49. Ross PT, Hart-Johnson T, Santen SA, et al. Considerations for using race and ethnicity as quantitative variables in medical education research. Perspect Med Educ. 2020 Oct;9(5):318–23.

50. Khine MS. Non-Cognitive Skills and Factors in Educational Success and Academic Achievement. In: Khine M.S., Areepattamannil S. (eds) Non-cognitive Skills and Factors in Educational Attainment. Contemporary Approaches to Research in learning Innovations. Rotterdam: SensePublishers; 2016. p. 3–9.

51. Gutman LM, Schoon I. The impact of non-cognitive skills on outcomes for young people. Institute of Education, London.; 2014.

52. MacKenzie RK, Dowell J, Ayansina D, et al. Do personality traits assessed on medical school admission predict exit performance? A UK-wide longitudinal cohort study. Adv Health Sci Educ. 2017 May;22(2):365–85.


All UK graduates attempting MRCS between April 2014 and May 2017
n= 3000
Excluded
n= 405
Matched Educational Performance Measure (EPM) data
n= 2590
Excluded
n= 5
Matched Situational Judgement Test (SJT) data
n= 2585
Matched MRCS Part A data
n= 2585
Matched MRCS Part B data
n= 755
Missing ethnicity data
n= 10
Excluded from adjusted regression analysis
Matched socio-demographic data
n= 2570
Matched socio-demographic data
n= 755


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Mean MRCS first-attempt pass rate
1st
2nd
3rd
4th
5th
6th
7th
8th
9th
10th
Educational Performance Measure Decile
Mean Part A Pass Rate
Mean Part B Pass Rate

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-7
Objectives	3	State specific objectives, including any prespecified hypotheses	4-7
Methods			
Study design	4	Present key elements of study design early in the paper	7-9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-9
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up (*b*) For matched studies, give matching criteria and number of exposed and unexposed	7-9
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-9
Bias	9	Describe any efforts to address potential sources of bias	7-9
Study size	10	Explain how the study size was arrived at	7-9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding (*b*) Describe any methods used to examine subgroups and interactions (*c*) Explain how missing data were addressed (*d*) If applicable, explain how loss to follow-up was addressed (*e*) Describe any sensitivity analyses	7-9
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	Figure 1 9-15
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	9-15
Outcome data	15*	Report numbers of outcome events or summary measures over time	9-15

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	9-15
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	9-15
Discussion			
Key results	18	Summarise key results with reference to study objectives	15-21
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15-21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-21
Generalisability	21	Discuss the generalisability (external validity) of the study results	15-21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Does performance at medical school predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination? A retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-046615.R2
Article Type:	Original research
Date Submitted by the Author:	06-Jul-2021
Complete List of Authors:	Ellis, Ricky; University of Aberdeen Institute of Applied Health Sciences; Nottingham University Hospitals NHS Trust, Urology Department Scrimgeour, Duncan; University of Aberdeen Institute of Applied Health Sciences; Aberdeen Royal Infirmary, Department of Colorectal Surgery Brennan, Peter ; Portsmouth Hospitals NHS Trust, Department of Oral and Maxillofacial Surgery Lee, Amanda; University of Aberdeen Institute of Applied Health Sciences, Department of Medical Statistics Cleland, Jennifer; Lee Kong Chian School of Medicine, Medical Education Research and Scholarship Unit (MERSU)
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Health policy, Surgery
Keywords:	MEDICAL EDUCATION & TRAINING, SURGERY, Adult surgery < SURGERY

Does performance at medical school predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination? A retrospective cohort study

Ricky Ellis[1,2,3], Duncan SG Scrimgeour[1,3], Peter A Brennan[4], Amanda J Lee[5], Jennifer Cleland[6]

[1]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen and [2]Urology Department, Nottingham University Hospitals, Nottingham and [3]Department of Colorectal Surgery, Aberdeen Royal Infirmary, Aberdeen and [4]Department of Maxillo-Facial Surgery, Queen Alexandra Hospital, Portsmouth and [5]Department of Medical Statistics, Institute of Applied Health Sciences, University of Aberdeen, [6]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.

Correspondence to: Mr Ricky Ellis

Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, AB25 2ZD

Rickyellis@nhs.net

ABSTRACT

Background: Identifying predictors of success in post-graduate examinations can help guide the career choices of medical students and may aid early identification of trainees requiring extra support to progress in specialty training. We assessed whether performance on the Educational Performance Measurement (EPM) and Situational Judgement Test (SJT) used for selection into Foundation Training predicted success at the Membership of the Royal College of Surgeons (MRCS) examination.

Methods: This was a longitudinal, cohort study using data from the UK Medical Education Database (https://www.ukmed.ac.uk). UK medical graduates who had attempted Part A (*n*=2585) and Part B (*n*=755) of the MRCS between 2014-2017 were included. Chi-squared and independent t-tests were used to examine the relationship between medical school performance and socio-demographic factors with first-attempt success at MRCS Part A and B. Multivariate logistic regression was employed to identify independent predictors of MRCS performance.

Results: The odds of passing MRCS increased by 55% for Part A (odds ratio 1.55 [95% confidence interval (CI) 1.48-1.61]) and 23% for Part B (1.23 [1.14-1.32]) for every additional EPM decile point gained. For every point awarded for additional degrees in the

EPM, candidates were 20% more likely to pass MRCS Part A (1.20 [1.13-1.29]) and 17% more likely to pass Part B (1.17 [1.04-1.33]). For every point awarded for publications in the EPM, candidates were 14% more likely to pass MRCS Part A (1.14 [1.01-1.28]). SJT score was not a statistically significant independent predictor of MRCS success.

Conclusion: This study has demonstrated the EPM's independent predictive power and found that medical school performance deciles are the most significant measure of predicting later success in the MRCS. These findings can be used by medical schools, training boards and workforce planners to inform evidence-based and contemporary selection and assessment strategies.

Key words: MRCS, UKFPO, situational judgement test, Medical School, Surgery, Postgraduate examinations

Strengths and limitations of this study

- This is the first study to investigate the relationship between medical school performance with performance at a high stakes UK postgraduate surgical examination.
- This is a large retrospective cohort study using the U.K. Medical Education database.
- This study examines whether performance on the Educational Performance Measurement (EPM) and Situational Judgement Test (SJT) used for selection into Foundation Training predicted success at the MRCS examination.
- Following previous studies, the relatively blunt measure of MRCS pass/fail results at first attempt was used as the primary outcome

INTRODUCTION

Progression through the United Kingdom (UK) medical education and training pathway is based on performance on a series of index assessments, starting with examination performance prior to entry to medical school and (typically) ending with respective Royal College Fellowship examinations. Each assessment is designed to ensure appropriate standards for stage of training and to ultimately safeguard patients. (1,2)

Performance at each stage also has implications on career progression. In the UK, doctors with higher academic scores during medical school are more likely to be offered their first choice of UK Foundation Programme (UKFP) training post on graduating. (3) Those with higher academic scores during medical school are also more likely to be offered a training place in a more competitive specialty. (4)

Studies have already demonstrated the validity of academic performance during medical school in predicting performance during Foundation Training,(5–7) although there is little research on the association between medical school performance and performance during specialty training in the UK. The seminal "Academic Backbone" paper by McManus et al.

(2013) described how prior attainment is the best predictor of future performance in medical education. (8) However, that study was carried out before standardized markers of medical school performance were introduced (see later) and therefore may not represent contemporary patterns of performance. This is an important deficiency in the literature as research from other contexts indicates that examination results taken during and shortly after medical school predict later performance on board certification examinations, and patient complaints. (9–14) Furthermore, if early assessments do not predict later performance, then their fitness for purpose as markers of performance and their use as gateways for progression in training are questionable.

At the time of this study the UK did not have a national licensing examination for graduating doctors. Instead, performance during medical school is measured within schools by the Educational Performance Measure, or EPM. (15) The EPM is calculated out of 50 points and comprises three parts (Table 1). The first component of the EPM is a quantitative measure of students' medical school performance compared to their peers (EPM Decile). Points are awarded depending on a student's final performance decile; ranging from 34 points for the 10^{th} (lowest) decile to 43 points for students in the 1^{st} (highest) decile. The EPM decile is calculated using multiple assessments of a student's knowledge and practical skills over time and by multiple assessments throughout medical school. (15) The second part of the EPM is comprised of points awarded for additional degrees (0–5 points are awarded according to degree grade achieved). The last part of the EPM is comprised of points that are awarded for publications (maximum 2 points; 1 point is awarded per publication). Points awarded for additional degrees and publications (0-7 in total) are described as Educational Achievements (EA). The selection process for the UKFP couples the EPM with a situational judgement test, or SJT (16–21) also scored out of 50 points. The UKFP SJT could be described as a type of 'procedural knowledge test'; assessing procedural knowledge about what to do in certain situations and how to do it. (21) The procedural knowledge being assessed by the UKFP SJT aligns with the behaviours and attitudes expected of doctors as described in the General Medical Council's (GMC) Good Medical Practice. (2) The graduate's combined EPM score plus SJT score out of 100 is their application score for the UKFP. (22)

Table 1. The components of the Educational Performance Measure used to quantify performance at medical school. The EPM is combined with a candidates SJT score to create their total UKFP selection score.

Educational Performance Measure (EPM)	Points available
Medical School Performance Decile	34-43
Additional Degrees	0-5
Publications	0-2
Maximum combined EPM score	**50**

Situational Judgment Test (SJT)	Points available
Test Score	50

United Kingdom Foundation Program (UKFP) Selection Score	Points available
EPM	50
SJT	50
Maximum UKFP Selection Score	**100**

The current study aimed to assess whether performance in medical school, EPM and SJT scores, could predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination. The MRCS examination is often taken by UK trainees during Foundation and Core surgical training years and comprises of two parts: Part A, a written examination with two papers and Part B, an objective structured clinical examination (OSCE). (23,24) The MRCS is a high-stakes postgraduate assessment that is used as a gateway for applications for higher surgical training and is itself a good predictor of future surgical training outcomes. (25–27)

Given performance at medical school and success in postgraduate assessments is related to socio-demographic factors as well as academic ability, regression models were adjusted for socio-demographic factors known to be associated with MRCS success. (3,23,25,28–30) These included gender, ethnicity and graduate status on entry to medical school. This analysis is timely given policy drivers in the UK to ensure that medical school and postgraduate assessments are fair,(1) the pending imposition of a once-off high-stakes test, the Medical Licensing Assessment (MLA),(31) and proposals to exclude Educational Achievements from the EPM score used in UKFP selection from 2023. (32)

Use of linked individual-level data from the UK Medical Education Database (UKMED: https://www.ukmed.ac.uk/) enabled a national-level analysis, drawing on data from sources including medical school assessment, Foundation Programme selection and postgraduate assessment outcomes. (33)

METHODS

A longitudinal retrospective cohort study was conducted on UK medical graduates who had attempted either the Part A (written) or the Part B (clinical) MRCS examination from April 2014 to May 2017.

The UK Medical Education Database (UKMED: https://www.ukmed.ac.uk/) was used to access linked data from UK medical schools and the four Royal Colleges of Surgeons in the UK and Ireland. All counts have been rounded according to Higher Education Statistics Agency (HESA) data standards to ensure person-level anonymity. (34)

The following data were extracted: self-declared gender, self-reported ethnicity demographics and graduation status at the time of entry to medical school, medical school Educational Performance Measure decile, additional degree and EPM publication scores, SJT score and MRCS Part A and B first attempt result. Figure 1 shows the flow of data through the study. Candidate first attempt results were used as they have been shown to be the best predictor of future performance in postgraduate examinations. (35)

Figure 1. Data flow through study.

Except for SJT and EPM scores, all variables were subsequently dichotomized. Graduation status was defined as "yes" if candidates had obtained a degree prior to entering medicine. Self-declared ethnicity was coded as "white" or "non-white" as used in similar studies to enable powered analysis of smaller cohorts. (25,26) Part A and B MRCS performance was categorized as "pass" 'or "fail" at first attempt.

Statistical analysis

All analyses were conducted using SPSS® v22.0 (IBM, Armonk, New York, USA). A chi-squared test was initially employed to determine any associations with first attempt MRCS pass/fail outcomes. The relationship between SJT, EPM decile, additional degrees, EPM publication scores and Part A and Part B MRCS first attempt success was examined using independent t-tests since the distribution of scores was normal. Correlation coefficients were calculated for Foundation Programme selection scores and the MRCS Part A (the written component of the MRCS examination) score relative to pass mark.

Logistic regression models were developed to identify predictors of success at MRCS at first attempt that were independent of other performance measures used in UKFP selection. Further regression models were developed to identify predictors of MRCS success, that were independent of other performance metrics and socio-demographic factors known to be associated with MRCS performance. While doctors are not selected for the UKFP based on socio-demographic factors, adjusting for these known predictors of MRCS success ensured that regression models were adjusted for these potential confounding factors and are therefore more applicable in real life. Potential interactions between significant predictors were also examined.

The Intercollegiate Committee for Basic Surgical Examinations (ICBSE) and its Internal Quality Assurance Subcommittee, which monitors standards and quality, approved this study. The highest standards of security, governance and confidentiality were ensured when storing, handling and analysing identifiable data.

Patient and Public involvement

No patients or public were involved in this study.

RESULTS

Part A MRCS

A total of 3000 UK medical graduates attempted Part A MRCS between April 2014 and May 2017. Of these 2585 had matched EPM and SJT data. 50% (n=1280) passed Part A MRCS at their first attempt. 63% of candidates (n=1635), were male, 56% were white (n=1435), and 81.5% had not undertaken a prior degree before entering medicine (n=2105). Mean (standard deviation [SD]) total EPM and SJT scores for candidates who had attempted Part A MRCS were 41.6 (3.86) and 39.4 (3.54) respectively.

Pass rates for Part A MRCS by gender, ethnicity and graduate on entry to medicine status are shown in Table 2. Differences in pass rates were statistically significant for: gender (54.9% males *vs.* 40.5% females, $P<0.001$), ethnicity (54.2% white *vs.* 44.1% non-white, $P<0.001$) and graduate status (50.7% no prior degree *vs.* 44.7% prior degree, $P=0.017$).

Table 2. Univariate analysis of MRCS first attempt pass rates by gender, ethnicity and graduation status for UK medical graduates.

Variable	Part A (*n*=2585)		Part B (*n*=755)	
	Pass *n*=1280	**Fail** *n*=1305	**Pass** *n*=575	**Fail** *n*=180
Gender				
Male	54.9% (895/1635)	45.1% (740/1635)	75.9% (380/500)	24.1% (120/500)
Female	40.5% (385/950)	59.5% (565/950)	77.1% (195/255)	22.9% (60/255)
p-value	< 0.001		0.719	
Ethnicity				
White	54.2% (780/1435)	45.8% (660/1435)	82.0% (350/430)	18.0% (75/430)
Non-white	44.1% (500/1135)	55.9% (635/1135)	68.8% (225/325)	31.2% (105/325)
Missing	*n*=10		*n*=0	
p-value	< 0.001		< 0.001	
Graduate on entry to medicine				
No	50.7% (1070/2105)	49.3% (1040/2105)	77.8% (495/635)	22.2% (140/635)
Yes	44.7% (210/480)	55.3% (265/480)	68.3% (80/120)	31.7% (40/120)
p-value	0.017		0.025	

Note. All p-values presented are from chi-squared analysis. MRCS, Membership of the Royal College of Surgeons.

Univariate analysis of EPM and SJT scores are shown in Table 3. Candidates who passed Part A MRCS at first attempt had performed better in their SJT (Mean 40.0 [SD 3.3] *vs.* 38.9 [3.7], $P<0.001$) and had scored higher for their total EPM (43.6 [3.3] *vs.* 39.8 [3.4], $P<0.001$) compared to those who failed at first attempt. Figure 2 shows the relative increase in mean MRCS Part A pass rates at first attempt according to candidate EPM decile.

Table 3. Univariate analysis of EPM scores, SJT scores and MRCS Part A and Part B first attempt success.

Variable		MRCS Part A Pass *n*=1280	MRCS Part A Fail *n*=1305	MRCS Part A P-value	MRCS Part B Pass *n*=575	MRCS Part B Fail *n*=180	MRCS Part B P-value
EPM Decile	Mean	40.1	37.3	<0.001	39.7	38.1	<0.001
	S.D.	2.4	2.5		2.6	2.5	
EPM Degree score	Mean	2.8	2.1	<0.001	2.7	2.3	0.001
	S.D.	1.5	1.6		1.5	1.6	
EPM Publications	Mean	0.7	0.5	<0.001	0.9	0.7	0.042
	S.D.	0.9	0.8		0.9	0.8	
SJT	Mean	40.0	38.9	<0.001	40.2	39.4	0.010
	S.D.	3.3	3.7		3.2	4.2	

Note. MRCS, Membership of the Royal College of Surgeons. EPM, Educational Performance Measure. SJT, Situational Judgment Test, N, number of candidates. S.D., Standard Deviation. *P*, *P* Value.

Figure 2. Relative increase in mean MRCS pass rates at first attempt according to candidate Educational Performance Measure (EPM) decile (1st EPM decile indicates the highest achieving candidates and 10th decile, the lowest achieving candidates).

Table 4 shows correlation coefficients between each Foundation Programme selection score and MRCS Part A. According to Cohen's guidelines (36) EPM degree score, EPM publication score and the SJT show statistically significant weak positive correlation with

Part A scores. Total EPM and EPM decile show statistically significant strong correlations with MRCS Part A.

Table 4. Correlation coefficients between Foundation Programme selection scores and MRCS Part A Scores (*n*= 2585).

	Correlation Coefficient	P-value
Total EPM	0.57	<0.001
EPM Decile*	0.59	<0.001
EPM Degree Score	0.27	<0.001
EPM Publications	0.17	<0.001
SJT	0.23	<0.001

*Spearman's Rho coefficient.

Table 5 shows the odds ratios (OR) and 95% confidence intervals (CI) for independent predictors of passing Part A MRCS at first attempt. Odds ratios were similar for UKFP selection metrics when multivariate analysis included socio-demographic predictors of MRCS success. EPM decile, EPM degree and EPM publication scores were predictors of MRCS success independent of other selection metrics and socio-demographic factors. Specifically, the odds of passing Part A MRCS at first attempt increased by 55% for every additional EPM decile (OR 1.55, 95% CI 1.48 to 1.61). The odds of passing Part A on first attempt increased by 20% for every additional EPM degree point (OR 1.20, 95% CI 1.13 to 1.29). Finally, the odds of passing Part A on first attempt increased by 14% for every additional EPM publication point awarded (OR 1.14, 95% CI 1.01 to 1.28). SJT score was not found to independently predict Part A first attempt success (*P*=0.177). There was a statistically significant interaction between ethnicity and gender in the final Part A MRCS regression model with white males more likely to pass, *P*=0.002. MRCS candidates who entered medical school without a prior degree were more than twice as likely to pass Part A compared to those who entered medical school as graduates (OR 2.23, 95% CI 1.73 to 2.87).

Table 5. Predictors of pass at first attempt at Part A and Part B MRCS for UK medical graduates on multivariate analysis. Odds ratios are unadjusted and then adjusted for socio-demographic predictors of MRCS success.

	Part A		Part B	
Variable	Unadjusted Odds ratio (95% CI) *n*= 2585	Adjusted Odds ratio (95% CI) *n*=2570	Unadjusted Odds Ratio (95% CI) *n*=755	Adjusted Odds Ratio (95% CI) *n*= 755
Educational Performance Measure Decile	1.50 (1.45 to 1.56)	1.55 (1.48 to 1.61)	1.23 (1.15 to 1.32)	1.23 (1.14 to 1.32)
Educational Performance Measure Additional Degree points	1.16 (1.09 to 1.23)	1.20 (1.13 to 1.29)	1.09 (0.97 to 1.23)	1.17 (1.04 to 1.33)

Educational Performance Measure	1.17	1.14	1.04	1.02
Publication Score	(1.04 to 1.31)	(1.01 to 1.28)	(0.85 to 1.28)	(0.83 to 1.26)
Situational Judgement Test	1.02	1.02	1.04	1.02
	(0.99 to 1.04)	(0.99 to 1.05)	(0.99 to 1.09)	(0.97 to 1.10)
Gender*	-	1.60	-	0.97
Males vs. Females		(1.17 to 2.10)		(0.66 to 1.43)
Ethnicity*	-	0.70	-	1.86
White vs. Non-White		(0.51 to 0.96)		(1.29 to 2.69)
Ethnicity*Gender	-	1.97	-	-
		(1.33 to 2.91)		
Graduate on entry into medicine	-	2.23	-	2.54
Non-Graduates vs. Graduates		(1.73 to 2.87)		(1.57 to 4.13)

MRCS, Membership of the Royal College of Surgeons; CI, Confidence Interval. *P=0.002 for interaction between Ethnicity and Gender in MRCS Part A Model.

Part B MRCS

In total, 755 of the Part A study cohort (n=2585) attempted MRCS Part B at a later date. 76.3% (n=575) of candidates passed Part B MRCS at first attempt. Unsurprisingly the demographics for Part B MRCS were similar to those observed for Part A MRCS candidates; 67% of candidates were male (n=500), 57% were white (n=430), and 84% had not undertaken a previous degree (n=635). The mean (SD) total EPM and SJT scores for UK graduates who had attempted Part B MRCS were 42.8 (3.67) and 40.0 (3.47) respectively.

Pass rates for Part B MRCS by gender, ethnicity and graduate on entry to medicine status are shown in Table 2. There was no significant difference in Part B MRCS first attempt pass rates between males and females (75.9% *vs.* 77.1% respectively P=0.719). Differences in pass rates were statistically significant for ethnicity (82.0% white *vs.* 68.8% non-white, P<0.001) and graduate status (77.8% no prior degree *vs.* 68.3% prior degree, P=0.025).

Univariate analysis of EPM and SJT scores are shown in Table 3. Those who passed Part B MRCS at first attempt had performed better in their SJT compared to those who failed at first attempt (40.2 [3.2] *vs.* 39.4 [4.2], P<0.010). Similarly, candidates who passed Part B at first attempt had scored higher in their total EPM (43.3 [3.6] *vs.* 41.1 [3.4], P<0.001). Figure 2 shows MRCS Part B performance according to EPM decile score. The overall trend reveals a relative increase in mean MRCS Part B pass rates at first attempt according to candidate EPM decile.

Table 5 shows the logistic regression models for independent predictors of Part B MRCS first attempt. EPM decile and EPM degree scores were statistically significant predictors of MRCS success independent of other selection metrics and socio-demographic factors. The odds of passing MRCS Part B at first attempt increased by 23% (OR 1.23, 95% CI 1.14 to 1.32) for every additional EPM decile. The odds of passing MRCS Part B at first attempt increased by 17% (OR 1.17, 95% CI 1.04 to 1.33) for every additional EPM degree point awarded. Neither SJT score or EPM publication scores were found to be independent predictors of Part B success at first attempt after adjusting for UKFP selection metrics and socio-demographic factors (*P*=0.429 and *P*=0.849 respectively).

White UK medical graduates were nearly twice as likely to pass Part B MRCS at first attempt compared to non-white candidates (OR 1.86, 95% CI 1.29 to 2.69). Candidates who had not undertaken a previous degree before entering medicine were more than twice as likely to pass Part B MRCS compared to those who had undertaken a prior degree (OR 2.54, 95% CI 1.57 to 4.13). There were no statistically significant interactions between any of the Part B MRCS variables in the adjusted regression model.

DISCUSSION

Educational Performance Measure - Deciles

We assessed the predictive validity of UKFP selection measures, the SJT and EPM, against the MRCS examination which is known to be a good predictor of future surgical training outcomes. (25–27) We found that EPM deciles predicted success at both Part A (written) and Part B (OSCE) of the MRCS independent of other UKFP selection scores and socio-demographic factors. For every incremental EPM decile, candidates were significantly more likely to pass both MRCS Part A and Part B. Reassuringly, the predictive value of EPM deciles were not significantly altered when adjusting for gender, ethnicity and graduate status, indicating that very little of the association that exists between FP selection scores and MRCS performance is explained by these socio-demographic factors. Our results add to previous studies which found that the EPM predicts performance during Foundation training,(6,7) and provides assurance that UK medical school assessments appropriately gauge student competence and readiness for practice.

A key limitation of the EPM decile score as a selection tool is its ranking of medical school graduates at a local rather than national level. Each medical school ranks their cohort of graduates internally into 10 equal groups (deciles) based on performance in a number of assessments taken over the duration of the medical course. Students are therefore ranked against their peers within each medical school, potentially penalising high-achieving individuals that study at more competitive schools, resulting in a lower decile score than if those individuals studied at schools with a less competitive cohort. Given that assessment also varies significantly in 'volume, type and intensity' between medical schools, concerns have been raised that students of equal proficiency may fall into different EPM deciles across schools due to differences in assessment rather than ability. (37) Furthermore, the number of assessments used and scoring for each assessment varies considerably between schools which can limit the range of scores used for decile ranking, reducing the spread of candidates. (38)

Concerns regarding the impact of variation in local assessment and ranking have resulted in demand for the MLA in the UK. A national MLA is argued to provide a potentially more robust method of ranking medical graduates nationally and may also contribute to standard setting for education across medical schools. The MLA's impending introduction has been met with a mixed response with some arguing that a single high-stakes exit examination is not as valuable as multiple local assessments over a number of years and may also result in schools teaching to pass instead of teaching to practice medicine. (37,39) However, a one-off high-stakes examination on completion of medical school reflects the use of assessments throughout postgraduate medical training. The predictive and incremental validity of the new MLA must be scrutinised to justify its financial cost and its burden on both students and the medical education system.

Despite the limitations and potential shortcomings of the EPM decile scoring system that is currently being used, it appears to achieve its intended function for UKFP selection. It differentiates between candidates by ability and demonstrates the ability to predict postgraduate performance. (6,40) These data support its predictive validity and ongoing use as a UKFP selection tool. How the EPM, a local-level assessment, will sit alongside the proposed MLA, remains to be seen.

Our findings also align with the "academic backbone" concept proposed by McManus et al. (2013); an idea that in medical education, current learning and achievement is dependent upon attainment at earlier stages. (8) This can be summarised simply as: medical students who are high achievers remain high achievers. Candidates ranked in the top deciles perform better at MRCS. Those who perform best in MRCS are more likely to achieve a specialty training (ST3) post at national selection, are more likely to progress through training with satisfactory Annual Review of Competence Progression (ARCP) outcomes and are more likely to succeed at first attempt in their surgical fellowship (FRCS) examinations. (25–27) The nature of the outcome measures have changed since the seminal study of McManus et al. in 2013,(8) but the principles have not: the road to success for those who wish to pursue a successful career in surgery begins early.

Educational Performance Measure - Educational Achievements (EAs)

Points awarded in the EPM for additional degrees predict success at MRCS independent of other UKFP selection measures and socio-demographic factors. While points awarded for additional publications independently predict success in MRCS Part A, they were not an independent predictor of success in MRCS Part B. Correlations between MRCS Part A scores and EA points were considerably weaker than the correlation with EPM decile scores. It also appears that EPM decile scores are largely responsible for the strength of the correlation seen between EPM total and MRCS Part A scores. These results are timely and relevant given the recent announcement that points awarded for EA will be excluded from the EPM scoring system for UKFP selection from 2023. (32)

Points awarded for EA in the EPM undoubtedly play a role in increasing the spread of applicant scores when combined with EPM decile and SJT points for UKFP selection. (40) However, there is evidence of increasing EA point inflation over recent years with the number of applicants earning EA points increasing from 30% to 70%. (41) Given this EA point inflation, it is possible that correlations between EA scores and MRCS performance found in our data may be higher than those seen in cohorts that have graduated from medical school more recently. It is clear that over time the ability of EA points to differentiate candidates will diminish, but the financial barriers to success in medicine that these may cause would persist, with students from more affluent backgrounds being in a position to 'pay for points' by studying an intercalated degree. Indeed, given the recent drive to widen access to medicine it would appear contradictory for selection tools to encourage students to take on the significant financial burden of an intercalated degree that is not necessary for the practice of medicine, and does not necessarily improve patient care. Studying an intercalated degree does undoubtedly have many advantages that would 'enrich the student experience',(42) but students should not be penalised in their national ranking if uninterested or unable to afford to do so.

Overall, it could be argued that the limited predictive value of EA points found in this study and others (6) does not outweigh their potential to limit the score and subsequent ranking of applicants' that are unable to afford to undertake an intercalated degree.

Foundation Programme Situational Judgement Test

Candidates who passed both parts A and B of the MRCS at first attempt scored higher in their SJT than candidates who failed, and there was a statistically significant positive correlation with Part A scores. However, the SJT did not independently predict MRCS success after adjusting for EPM scores and socio-demographic factors, displaying no significant incremental value over and above the predictive value of EPM decile scores. It is important to consider this finding in relation to the premise behind SJTs.

The Foundation Programme (FP) SJT is based on a job analysis of being a Foundation Doctor. (7) Significant correlation between SJT and EPM scores between schools has been identified. (6,37,43) Additionally both SJT and EPM scores are independently associated with the odds of successful completion of the FP, and SJT score offers a degree of incremental predictive validity over that provided by the EPM deciles, suggesting that it is capturing additional, relevant, information on applicants, as intended. (6,40) Research suggests that the FP SJT does what it was designed to do as well as succeeding in increasing the spread of candidates being ranked for foundation training posts (the arguments as to how it should be weighed in the FP selection process are outside the scope of this paper but we direct readers with an interest to other studies). (6,44,45) It was not designed to select for specialty training: where specialty training programmes use SJTs for selection, these have been designed specifically against the role of a trainee/resident in that speciality. (46–48) Given this, in retrospect it is unsurprising that the FP SJT does not independently predict performance on a post-graduate examination that tests the clinical knowledge, skills and professional attitudes expected of surgical trainees.

Strengths and Weaknesses

The current study is one of the first to use the UKMED to examine the associations between medical school performance and FP SJT outcomes on success at a high-stakes postgraduate surgical examination. The UKMED enabled a nationwide, multi-cohort analysis and our breakdown of the FP selection process into EPM scores and SJT allowed us to look separately at academic attainment and other factors.

There are some limitations of the study. Firstly, although candidates can take Parts A and B MRCS on multiple occasions, we used candidate first attempt results as the best predictor of future performance. (35) We often used the relatively blunt outcome measure of pass/fail as this is what is meaningful to those sitting the MRCS, and has been used in previous studies looking at factors that predict performance in the MRCS. (25) Self-declared ethnicity data were combined into two discrete categories to maximise power when analysing smaller cohorts, rather than this being an ethical or social decision. Regression analyses were adjusted for known socio-demographic predictors of MRCS success, but these were not the main focus of the current paper. We are currently undertaking further analyses to characterise group-level attainment differences that have been identified. (49) Finally, the current analysis was based on retrospective quantitative data. A prospective study would have allowed us to examine more variables related to progression and attainment in surgical training. For example, being good at passing exams is linked to academic ability, but the wider education literature makes clear that non-cognitive factors such as motivation, time management and resilience are also relevant to performance. (50,51) If appropriate measures could be identified,(52) it would be interesting to compare graduates and MRCS candidates on these factors.

CONCLUSION

Success at first attempt of MRCS Part A and B can be predicted from medical school performance (EPM decile score) but not from the Foundation Programme SJT score. Put simply, medical students who do well in terms of medical school examination performance remain strong performers later on in their careers. These results may help to guide career choices for students and can be used by training institutions to inform evidence-based and contemporary selection and assessment strategies.

ABBREVIATIONS

ARCP: Annual Review of Competence Progression

EA: Educational Achievements

EPM: Educational Performance Measure

FRCS: Fellowship of the Royal College of Surgeons Examinations

GMC: General Medical Council

HESA: Higher Education Statistics Agency

MLA: United Kingdom Medical Licensing Examination

MRCS: Intercollegiate Membership of the Royal College of Surgeons Examinations

SJT: Situational Judgement Test

UKFP: United Kingdom Foundation Program

USMLE: United States Medical Licensing Examinations

ACKNOWLEDGEMENTS

The authors would like to acknowledge Iain Targett at the Royal College of Surgeons of England, for his help with data collection and Gregory Ayre from the Intercollegiate Committee for Basic Surgical Examinations for their support during this project. Our thanks to members of the UKMED Research Group who provided useful feedback on an earlier version of this manuscript, and whose comments helped refine the paper. The authors would also like to acknowledge Daniel Smith for his help with the UKMED database. Data Source: UK Medical Education Database ("UKMED"). UKMEDP043 extract generated on 25/07/2018. We are grateful to UKMED for the use of these data. However, UKMED bears no responsibility for their analysis or interpretation. The data includes information derived from that collected by the Higher Education Statistics Agency Limited ("HESA") and provided to the GMC ("HESA Data"). Source: HESA Student Records 2007/2008 to 2015/2016. Copyright Higher Education Statistics Agency Limited. The Higher Education Statistics Agency Limited makes no warranty as to the accuracy of the HESA Data, cannot accept responsibility for any inferences or conclusions derived by third parties from data or other Information supplied by it.

FUNDING

Royal College of Surgeons of England, Royal College of Surgeons of Edinburgh, Royal College of Surgeons of Ireland and Royal College of Physicians and Surgeons of Glasgow (Award/Grant number is not applicable).

AVAILABILITY OF DATA AND MATERIALS

The dataset used in this study was acquired from the UK Medical Education Database and is held in Safe Haven. Data access requests must be made to UKMED. Full information for applications can be found at https://www.ukmed.ac.uk.

CONFLICTS OF INTEREST

None to declare.

ETHICAL APPROVAL

No formal ethical approval was required for this study of existing UKMED data. UKMED has received ethics exemption for projects using exclusively UKMED data from Queen Marys University of London Ethics of Research Committee on behalf of all UK medical schools (https://www.ukmed.ac.uk/documents/UKMED_research_projects_ethics_exemption.pdf). The Intercollegiate Committee for Basic Surgical Examinations (ICBSE) and its Internal Quality Assurance Subcommittee, which monitors MRCS standards, research and quality, approved this study.

CONTRIBUTORSHIP

RE and DS wrote the first draft of the manuscript. RE and DS performed statistical analyses with AL's supervision. RE, DS, PB, AL and JC reviewed and edited the manuscript. JC led the study proposal for access to UKMED data. All authors approved final draft of the manuscript.

REFERENCES

1. Promoting excellence: standards for medical education and training [Internet]. General Medical Council; 2015 [cited 2021 May 21]. Available from: www.gmc-uk.org/education/standards.asp

2. Good Medical Practice [Internet]. General Medical Council; 2013 [cited 2021 May 21]. Available from: www.gmc-uk.org/guidance.

3. Kumwenda B, Cleland JA, Prescott GJ, et al. Relationship between sociodemographic factors and selection into UK postgraduate medical training programmes: a national cohort study. BMJ Open. 2018 Jun;8(6):e021329.

4. Kumwenda B, Cleland J, Prescott G, et al. Relationship between sociodemographic factors and specialty destination of UK trainee doctors: a national cohort study. BMJ Open. 2019 Mar;9(3):e026961.

5. Patterson F, Ashworth V, Murray H, et al. Analysis of the Situational Judgement Test for Selection to the Foundation Programme 2013. Technical Report. [Internet]. Work Psychology Group; 2013 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2017/04/fy1-sjt-2013-technical-report.pdf

6. Smith DT, Tiffin PA. Evaluating the validity of the selection measures used for the UK's foundation medical training programme: a national cohort study. BMJ Open. 2018 Jul 1;8(7):e021918.

7. Medical Schools Council on behalf of the cross-stakeholder Project Group. Improving Selection to the Foundation Programme Final Report of the Parallel Recruitment Exercise. [Internet]. 2011 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2017/04/isfp-final-report-appendices.pdf

8. McManus I, Woolf K, Dacre J, et al. 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 Dec;11(1):242.

9. de Virgilio C, Yaghoubian A, Kaji A, et al. Predicting performance on the American Board of Surgery qualifying and certifying examinations: a multi-institutional study. Arch Surg Chic Ill 1960. 2010 Sep;145(9):852–6.

10. Tamblyn R, Abrahamowicz M, Brailovsky C, et al. Association between licensing examination scores and resource use and quality of care in primary care practice. JAMA. 1998 Sep 16;280(11):989–96.

11. Tamblyn R, Abrahamowicz M, Dauphinee D, et al. Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. JAMA. 2007 Sep 5;298(9):993–1001.

12. McCaskill QE, Kirk JJ, Barata DM, et al. USMLE step 1 scores as a significant predictor of future board passage in pediatrics. Ambul Pediatr Off J Ambul Pediatr Assoc. 2007 Apr;7(2):192–5.

13. Wenghofer E, Klass D, Abrahamowicz M, et al. Doctor scores on national qualifying examinations predict quality of care in future practice. Med Educ. 2009 Dec;43(12):1166–73.

14. Norcini JJ, Boulet JR, Opalek A, et al. The relationship between licensing examination performance and the outcomes of care by international medical school graduates. Acad Med J Assoc Am Med Coll. 2014 Aug;89(8):1157–62.

15. Educational Performance Measure (EPM) 2020 Framework [Internet]. UK Foundation Programme; 2020 [cited 2021 May 21]. Available from: https://foundationprogramme.nhs.uk/resources/

16. Patterson F, Driver R. Situational Judgement Tests (SJTs). In: Patterson F, Zibarras L, editors. Selection and Recruitment in the Healthcare Professions [Internet]. Springer International Publishing; 2018 [cited 2021 May 21]. p. 79–112. Available from: http://link.springer.com/10.1007/978-3-319-94971-0_4

17. Patterson F, Knight A, Dowell J, et al. How effective are selection methods in medical education? A systematic review. Med Educ. 2016 Jan;50(1):36–60.

18. Mcdaniel M, Morgeson F, Finnegan E, et al. Use of Situational Judgment Tests to Predict Job Performance. J Appl Psychol. 2001 Sep 1;86:730–40.

19. McDaniel MA, Hartman NS, Whetzel DL, et al. Situational Judgment Tests, Response instructions and Validity: A Meta-Analysis. Pers Psychol. 2007 Mar;60(1):63–91.

20. Lievens F, Peeters H, Schollaert E. Situational judgment tests: a review of recent research. Pers Rev. 2008 Jun 6;37(4):426–41.

21. Webster ES, Paton LW, Crampton PES, Tiffin PA. Situational judgement test validity for selection: A systematic review and meta-analysis. Med Educ. 2020 Oct;54(10):888–902.

22. UK Foundation Programme 2020 Applicants Handbook. [Internet]. 2020 [cited 2021 May 21]. Available from: https://www.foundationprogramme.nhs.uk/wp-content/uploads/sites/2/2019/10/UKFP-2020-Applicants-Handbook-.pdf

23. Scrimgeour DSG, Cleland J, Lee AJ, et al. Which factors predict success in the mandatory UK postgraduate surgical exam: The Intercollegiate Membership of the Royal College of Surgeons (MRCS)? Surg J R Coll Surg Edinb Irel. 2018 Aug;16(4):220–6.

24. Scrimgeour D, Cleland J, Lee A, Brennan PA. When is the best time to sit the MRCS examination? BMJ. 2017 Feb 6;356:j461.

25. Scrimgeour DSG, Cleland J, Lee AJ, et al. Prediction of success at UK Specialty Board Examinations using the mandatory postgraduate UK surgical examination. BJS Open. 2019 Dec;3(6):865–71.

26. Scrimgeour D, Brennan P, Griffiths G, et al. Does the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination predict 'on-the-job' performance

during UK higher specialty surgical training? Ann R Coll Surg Engl. 2018 Nov;100(8):669–75.

27. Scrimgeour DSG, Cleland J, Lee AJ, et al. Impact of performance in a mandatory postgraduate surgical examination on selection into specialty training: BJS Open. 2017 Oct;1(3):67–74.

28. Richens D, Graham TR, James J, et al. Racial and Gender Influences on Pass Rates for the UK and Ireland Specialty Board Examinations. J Surg Educ. 2016 Jan 1;73(1):143–50.

29. Puddey IB, Mercer A, Carr SE. Relative progress and academic performance of graduate vs undergraduate entrants to an Australian medical school. BMC Med Educ. 2019 May 22;19(1):159.

30. Woolf K, Cave J, Greenhalgh T, et al. Ethnic stereotypes and the underachievement of UK medical students from ethnic minorities: qualitative study. BMJ. 2008;337(a1220).

31. Medical Licensing Assessment [Internet]. General Medical Council; 2020 [cited 2021 May 21]. Available from: https://www.gmc-uk.org/education/medical-licensing-assessment

32. For UKFP2023 Entry – Update on Educational Achievements [Internet]. UK Foundation Programme; 2020 [cited 2021 May 21]. Available from: https://healtheducationengland.sharepoint.com/:w:/s/UKFPOT/EYd3TEAU749MpYfS4tWDEO8Bx-kdOsDmPs73uTDAH-xzxA?e=d3HdRi

33. Dowell J, Cleland J, Fitzpatrick S, et al. The UK medical education database (UKMED) what is it? Why and how might you use it? BMC Med Educ. 2018 Dec;18(1):6.

34. Rounding and suppression to anonymise statistics | HESA [Internet]. [cited 2021 May 21]. Available from: https://www.hesa.ac.uk/about/regulation/data-protection/rounding-and-suppression-anonymise-statistics

35. McManus I, Ludka K. Resitting a high-stakes postgraduate medical examination on multiple occasions: nonlinear multilevel modelling of performance in the MRCP(UK) examinations. BMC Med. 2012 Jun 14;10(1):60.

36. Hemphill JF. Interpreting the magnitudes of correlation coefficients. Am Psychol. 2003;58(1):78–9.

37. Devine OP, Harborne AC, McManus IC. Assessment at UK medical schools varies substantially in volume, type and intensity and correlates with postgraduate attainment. BMC Med Educ. 2015 Sep 11;15(1):146.

38. Improving Selection to the Foundation Programme - Final Report [Internet]. Medical Schools Council on behalf of the cross-stakeholder Project Group; 2011 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2017/04/isfp-final-report-appendices.pdf

39. Noble ISG. Are national qualifying examinations a fair way to rank medical students? No. BMJ. 2008 Aug 22;337:a1279.

40. Patterson F, Kerrin M, Edwards H, et al. Validation of the F1 Selection Tools [Internet]. Work Psychology Group on behalf of Health Education England; 2015 [cited 2021 May 21]. Available from: https://isfporguk.files.wordpress.com/2018/06/validation_of_the_f1_selection_tools_report_final_for_publication-1.pdf

41. For UKFP2023 Entry – Update on Educational Achievements (December 2020 Follow-Up) [Internet]. UK Foundation Programme; 2020 [cited 2021 May 21]. Available from: https://healtheducationengland.sharepoint.com/:b:/g/UKFPO/EYiV65BOKgpJsu4qLnakqPIBNdNM9HQ7vl4GMMALeCP4Dg?e=ZTVjgY

42. MSC Statement on the UKFPO decision to remove Educational Achievements from the Foundation ranking process [Internet]. Medical Schools Council [cited 2021 May 21]. Available from: https://www.medschools.ac.uk/news/msc-statement-on-the-ukfpo-decision-to-remove-educational-achievements-from-the-foundation-ranking-process

43. McManus IC, Harborne AC, Horsfall HL, et al. Exploring UK medical school differences: the MedDifs study of selection, teaching, student and F1 perceptions, postgraduate outcomes and fitness to practise. BMC Med. 2020 May 14;18(1):136.

44. Najim M, Rabee R, Sherwani Y, et al. The situational judgement test: a student's worst nightmare. Adv Med Educ Pract. 2015 Oct 5;6:577–8.

45. Mclachlan JC, Illing J. An Independant review of events relating to the use of Situational Judgement Tests in the selection for Foundation process 2013. Commissioned by Health Education England [Internet]. 2013 [cited 2021 May 21]. Available from: https://www.medschools.ac.uk/media/1932/hee-independent-review-of-the-sjt.pdf

46. Patterson F, Rowett E, Hale R, et al. The predictive validity of a situational judgement test and multiple-mini interview for entry into postgraduate training in Australia. BMC Med Educ. 2016 Dec;16(1):87.

47. Applicant Guidance. General Practice: Multi-Specialty Recruitment Assessment (MSRA) [Internet]. Health Education England; [cited 2020 May 19]. Available from: https://gprecruitment.hee.nhs.uk/Resource-Bank/Recruitment-Documents-Forms

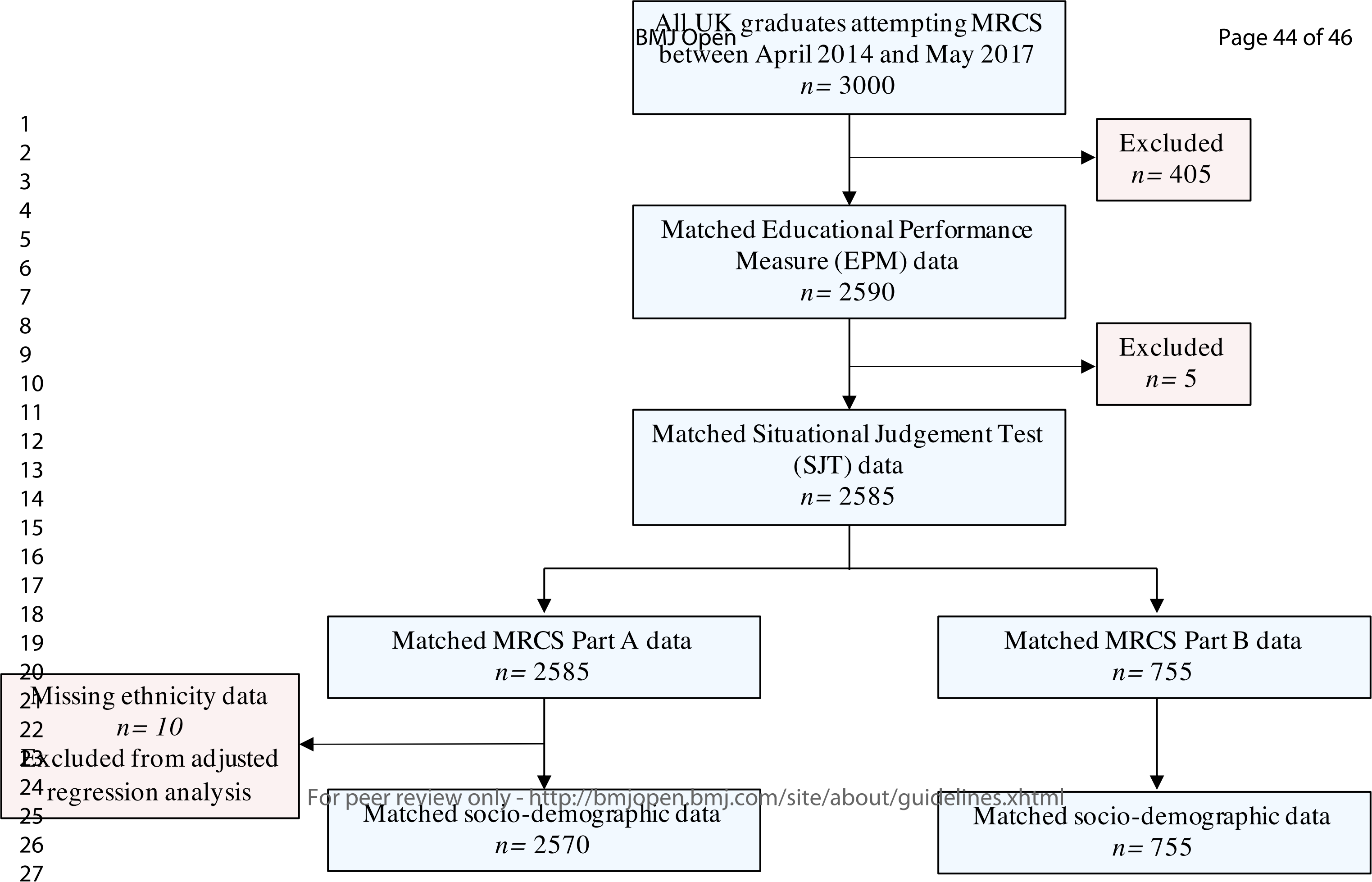
48. Multi-Specialty Recruitment Assessment (MSRA) – Test Blueprint & Information [Internet]. Work Psychology Group on behalf of Health Education England; 2019 [cited 2021 May 21]. Available from: https://gprecruitment.hee.nhs.uk/Resource-Bank/Recruitment-Documents-Forms

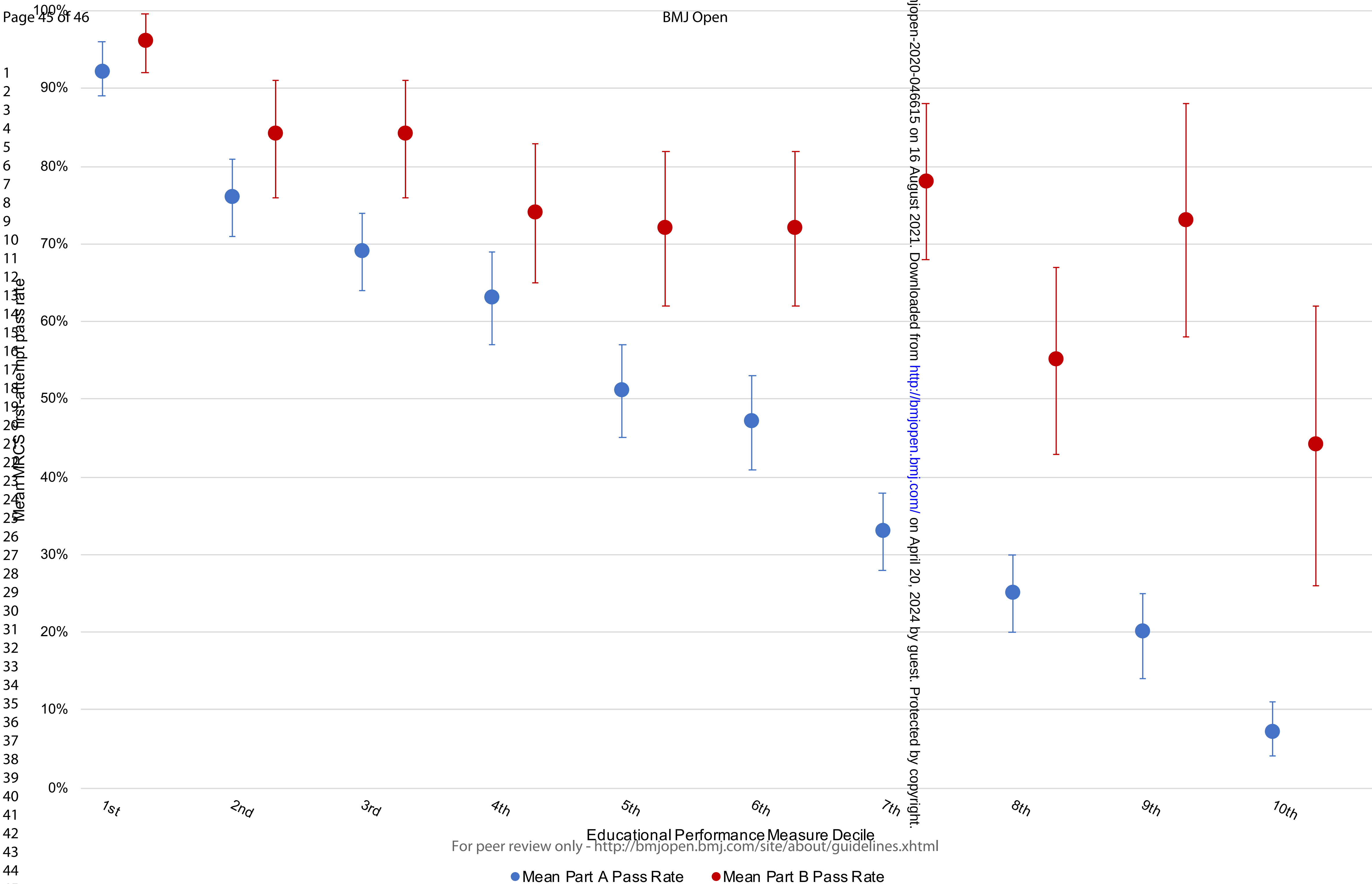
49. Ross PT, Hart-Johnson T, Santen SA, et al. Considerations for using race and ethnicity as quantitative variables in medical education research. Perspect Med Educ. 2020 Oct;9(5):318–23.

50. Khine MS. Non-Cognitive Skills and Factors in Educational Success and Academic Achievement. In: Khine M.S., Areepattamannil S. (eds) Non-cognitive Skills and Factors in Educational Attainment. Contemporary Approaches to Research in learning Innovations. Rotterdam: SensePublishers; 2016. p. 3–9.

51. Gutman LM, Schoon I. The impact of non-cognitive skills on outcomes for young people. Institute of Education, London.; 2014.

52. MacKenzie RK, Dowell J, Ayansina D, et al. Do personality traits assessed on medical school admission predict exit performance? A UK-wide longitudinal cohort study. Adv Health Sci Educ. 2017 May;22(2):365–85.


All UK graduates attempting MRCS between April 2014 and May 2017
n= 3000
Excluded
n= 405
Matched Educational Performance Measure (EPM) data
n= 2590
Excluded
n= 5
Matched Situational Judgement Test (SJT) data
n= 2585
Matched MRCS Part A data
n= 2585
Matched MRCS Part B data
n= 755
Missing ethnicity data
n= 10
Excluded from adjusted regression analysis
Matched socio-demographic data
n= 2570
Matched socio-demographic data
n= 755


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Mean MRCS first-attempt pass rate
1st
2nd
3rd
4th
5th
6th
7th
8th
9th
10th
Educational Performance Measure Decile
Mean Part A Pass Rate
Mean Part B Pass Rate

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-7
Objectives	3	State specific objectives, including any prespecified hypotheses	4-7
Methods			
Study design	4	Present key elements of study design early in the paper	7-9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-9
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up (*b*) For matched studies, give matching criteria and number of exposed and unexposed	7-9
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-9
Bias	9	Describe any efforts to address potential sources of bias	7-9
Study size	10	Explain how the study size was arrived at	7-9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding (*b*) Describe any methods used to examine subgroups and interactions (*c*) Explain how missing data were addressed (*d*) If applicable, explain how loss to follow-up was addressed (*e*) Describe any sensitivity analyses	7-9
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	Figure 1 9-15
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	9-15
Outcome data	15*	Report numbers of outcome events or summary measures over time	9-15

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	9-15
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	9-15
Discussion			
Key results	18	Summarise key results with reference to study objectives	15-21
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15-21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-21
Generalisability	21	Discuss the generalisability (external validity) of the study results	15-21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.