Establishing the predictive validity of the intercollegiate membership of the Royal Colleges of surgeons written examination: MRCS Part A

Ricky Ellis [a,b,*], Jennifer Cleland [c], Duncan SG. Scrimgeour [a,d], Amanda J. Lee [e], John Hines [f], Peter A. Brennan [g]

[a] Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, AB25 2ZD, United Kingdom
[b] Urology Department, Nottingham University Hospitals, Nottingham, United Kingdom
[c] Medical Education Research and Scholarship Unit, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
[d] Department of Colorectal Surgery, Aberdeen Royal Infirmary, Aberdeen, AB25 2ZN, United Kingdom
[e] Medical Statistics Team, Institute of Applied Health Sciences, University of Aberdeen, AB25 2ZD, United Kingdom
[f] Urology Department, University College Hospital, London, W1G 8PH, United Kingdom
[g] Department of Maxillo-Facial Surgery, Queen Alexandra Hospital, Portsmouth, PO6 3LY, United Kingdom

ARTICLE INFO

Article history:
Received 14 April 2023
Received in revised form
12 July 2023
Accepted 13 July 2023
Available online 4 August 2023

Keywords:
MRCS
Postgraduate
Validity
Training
Education

ABSTRACT

Successful completion of the Intercollegiate Membership of the Royal Colleges of Surgeons (MRCS) examination is mandatory for surgical trainees entering higher specialist training in the United Kingdom. Despite its international reputation, and the value placed on the examination in surgical training, there has been little evidence of its predictive validity until recently. In this review, we present a summary of findings of four recent Intercollegiate studies assessing the predictive validity of the MRCS Part A (written) examination.

Data from all four studies showed statistically significant positive correlations between the MRCS Part A and other written examinations taken by surgical trainees over the course of their education. The studies summarised in this review provide compelling evidence for the predictive validity of this gatekeeping examination. This review will be of interest to trainees, training institutions and the Royal Colleges given the value placed on the examination by surgical training programmes.

Introduction

The Intercollegiate Membership of the Royal Colleges of Surgeons (MRCS) examination is a high-stakes postgraduate assessment taken by more than 6000 surgical trainees every year. The MRCS acts as a safeguard for patients ensuring that specialist surgical trainees have met a universally respected standard. In the UK, success at MRCS is an indicator that trainees have acquired the knowledge, skills, attitudes and attributes expected of them at the completion of Core Surgical Training (CST). As such, it is highly valued as a gatekeeper to the surgical profession.[1,2] Completion of both Parts A (written) and B (Objective Structured Clinical Examination (OSCE)) is mandatory for surgical trainees applying for higher specialist training (HST) programmes in the United Kingdom (UK).

The examination is taken at considerable personal, social and financial cost to trainees,[3] and failure can have significant implications for career progression. Given the burden that this assessment places on trainees, it is vital that the examination is reliable and valid. If the examination is to continue to be used as a benchmark for surgical trainees worldwide, we must

* *Corresponding author*. Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, AB25 2ZD, United Kingdom.
E-mail addresses: Rickyellis@nhs.net (R. Ellis), jennifer.cleland@ntu.edu.sg (J. Cleland), duncan.scrimgeour@nhs.scot (D.SG. Scrimgeour), a.j.lee@abdn.ac.uk (A.J. Lee), Jhines@rcseng.ac.uk (J. Hines), peter.brennan@porthosp.nhs.uk (P.A. Brennan).
https://doi.org/10.1016/j.surge.2023.07.004

first ensure that it is fit for purpose and that it achieves its aims and objectives as a gatekeeper for surgical training.

Until recently, little was known about the predictive validity of the MRCS examination. This article discusses the principles of examination reliability and validity. It also summarises the findings of recent studies undertaken to investigate the predictive validity of the MRCS Part A written examination.

MRCS Part A

MRCS comprises two parts; Part A, the written component and Part B, the clinical examination. Part A includes a 3-h single best answer multiple-choice questionnaire (MCQ) paper assessing 'Applied Basic Sciences' followed by a 2-h MCQ paper assessing the 'Principles of Surgery in General'.[4] A minimum level of competence must be reached in each of the papers in addition to achieving the overall combined pass mark for both papers (as determined using a modified Angoff method, commonly used by other institutions delivering postgraduate examinations) to pass MRCS Part A. Candidates must pass Part A to be eligible to attempt MRCS Part B.

Reliability

A requirement for quality assurance for any postgraduate examination is the ability to demonstrate its reliability and validity.[5] The Intercollegiate Committee for Basic Surgical Examinations (ICBSE) was created by the Joint Royal Colleges of the UK and Ireland to develop, maintain and quality assure both the MRCS and Ear Nose and Throat (ENT) Diploma in Otolaryngology and Head and Neck Surgery (DOHNS) examinations. Reliability is a measure of the reproducibility of the examination and its results.[6,7] The ICBSE conducts rigorous testing of examination reliability annually, including measurement error, inter-rater, and test-retest reliability. Measurements of internal consistency are published open-access[8] and range from 0.95 to 0.96 (using Kuder-Richardson formula 20[9]) for the MRCS Part A and 0.59–0.88 (using Cronbach's alpha) for Part B. These results are comparable to those of other national and international postgraduate medical examinations.[10]

Validity

The validity of an examination describes whether it is measuring what it intends to (i.e. it is meeting its objective). It is more difficult to assess than reliability, requiring the comparison of multiple sources of assessment.[6,11] It is widely accepted that for a medical examination to be valid, it must demonstrate face validity (ensuring that the examination tests what it intends to test), content validity (the extent to which the examination tests knowledge of the curriculum) and predictive validity (the ability of a test to predict future outcomes).

With regards to how these validity measures are applied to the MRCS; the *content* of the MRCS examination is mapped to both the Intercollegiate Surgical Curriculum Programme and the General Medical Council (GMC) framework on Generic Professional Capabilities and is published in a guide for candidates.[4] Stringent correlation with these curricula ensures the *face validity* of the examination whilst questions and OSCE stations are drawn from a large question bank but are carefully regulated to ensure it examines the entire curriculum fairly to maintain *content validity*.

Predictive validity

As well as measures of internal validity, the examination should demonstrate its ability to predict future outcomes. However, until recently, MRCS predictive validity has remained largely untested, unlike other UK postgraduate medical examinations such as the Membership of the Royal College of Physicians (MRCP), the Membership of the Royal College of General Practitioners (MRCGP), Professional and Linguistic Assessment Board (PLAB) test and overseas including the American Board of Surgery qualifying and certifying examinations, United States Medical Licensing Examination (USMLE), and the Canadian Licensing Examinations.[12–17] This is an important gap in the literature as research has shown performance on other postgraduate medical examinations to predict later clinical performance and patient complaints.[18–21]

In the absence of a gold standard with which to compare medical examinations, the GMC confirmed that one way of validating an assessment method is to 'establish the strength of the relationships between similar assessments'.[22] Performance in one test should predict the performance in a future similar test (*predictive validity*) and correlate with performance in previous tests. If early assessments do not predict later success, then their fitness for purpose as markers of performance and their use as gateways for progression in training is questionable. The association between assessment outcomes may also be compared between MRCS and other examinations taken earlier in trainees' educational careers with the expectation that these should demonstrate significant linear correlations.

Background to the UK surgical training pathway

Figure 1 is a graphical representation of surgical training undertaken by most UK trainees. This figure also includes most of the key assessments that surgical trainees are likely to have undertaken during their training that can be used to evaluate the predictive validity of MRCS Part A.

Medical school selection in the UK includes three stages. The first is the use of school-exit examination performance (A-Levels or their equivalent, e.g., Scottish Highers, Irish Leaving certificates or International Baccalaureate) as a marker of prior academic achievement. The second stage is usually performance on one of the following selection tests: the University Clinical Aptitude Test (UCAT), Biomedical Admissions Test (BMAT) or Graduate Medical School Admissions Test (GAMSAT). Those identified as having the potential to study medicine from these two stages are usually invited to interview. This process is described in detail by Cleland et al., 2016 and the Medical Schools Council.[23,24]

Performance on completion of UK medical school is quantified by the Educational Performance Measure (EPM), a score comprised of; a student's performance decile within

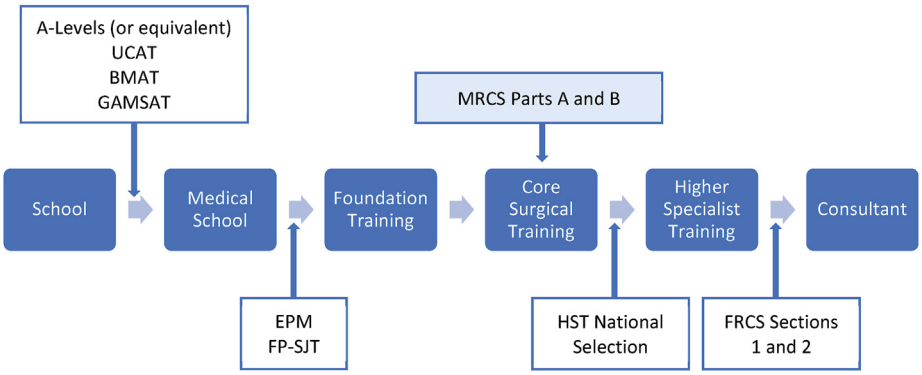

Figure 1 – UK Surgical training pathway, including assessments taken throughout. Please see the list of abbreviations.

each medical school, with additional points awarded for peer-reviewed publications and previous degree-level qualifications.[25] For selection into the UK Foundation Training Programme (FP), medical school graduates also sit a Situational Judgement Test (FP-SJT).[26] Although not a measure of performance at medical school, the FP-SJT is a written examination that aims to test the behaviours, traits and attitudes expected of doctors as described in the GMC's Good Medical Practice.[27] At the time of writing this paper, a combined score of the EPM and FP-SJT is used to rank each graduate nationally for allocation to FP Training posts.[28]

Lastly, the FRCS examination, taken during HST, is comprised of Section 1, a written examination with two papers, and Section 2, an OSCE examination.[29] Successful completion of both parts is a prerequisite for the award of Certificate of Completion of Training (CCT), enabling a surgeon to apply for Consultant posts.[30]

Methods

ICBSE established a research fellowship programme with the aim of investigating the predictive validity of the MRCS examination.[31] Several large longitudinal cohort studies have been undertaken by the Fellows and their research teams. In this review, we summarise the findings of these studies and contextualise data in establishing the predictive validity of the MRCS Part A examination.

Four recent longitudinal cohort studies by the ICBSE research group are presented with their combined utility in the context of assessing the predictive validity of the MRCS Part A examination. A description of each study and its primary outcomes is shown in Table 1. All four studies used a combination of univariate analyses, Pearson correlation coefficients and logistic regression modelling to assess the relationship between each examination and the MRCS Part A.

Table 1 – Primary outcomes measured in each of the four studies by the ICBSE Research group.

Study	Title	Assessments Compared to MRCS Part A	Number of Candidates
Ellis et al., 2021 *Postgraduate Medical Journal*.[41]	Performance at medical school selection correlates with success in Part A of the Intercollegiate Membership of the Royal Colleges of Surgeons (MRCS) Examination	A-Levels University Clinical Aptitude Test (UCAT) Biomedical Admissions Test (BMAT) Graduate Medical School Admissions Test (GAMSAT)	11,570
Ellis et al., 2021 BMJ *Open*.[42]	Does performance at medical school predict success at the Intercollegiate Membership of the Royal Colleges of Surgeons (MRCS) examination? A retrospective cohort study	Educational Performance Measure (EPM) Decile EPM Publication Score EPM Degree Score Foundation Programme Situational Judgement Test (FP-SJT)	2585
Scrimgeour et al., 2018 *The Surgeon*.[43]	Which factors predict success in the mandatory UK postgraduate surgical exam: The Intercollegiate Membership of the Royal Colleges of Surgeons (MRCS)?	Membership of the Royal Colleges of Surgeons (MRCS) Part B	7896
Scrimgeour et al., 2019 *BJS Open*.[44]	Prediction of success at UK Specialty Board Examinations using the mandatory postgraduate UK surgical examination	Fellowship of the Royal Colleges of Surgeons (FRCS) Section 1	854

All studies used MRCS results at the first attempt as this has been shown to be the best predictor of future performance in postgraduate examinations.[32]

Results

Table 2 shows the Pearson correlation coefficients between each testing method analysed across all four papers and MRCS Part A first-attempt scores. The table also displays the number of candidates included in each study.[33–36]

Except for the written communication subtest of the GAMSAT examination, a statistically significant correlation was found between all medical school selection test scores and MRCS Part A scores at the first attempt ($p < 0.001$).[33] According to Cohen's guidelines[37] weak positive correlation was found between MRCS Part A and both A-level ($r = 0.17–0.22$) and UCAT scores ($r = 0.25–0.26$). Significant moderate correlations were found between Part A and BMAT ($r = 0.29–0.33$) and GAMSAT ($r = 0.38$) scores. Total A-Level scores, UCAT, BMAT and GAMSAT scores were all significantly higher for candidates who passed MRCS Part A at the first attempt compared to those who failed on first attempt ($p < 0.05$). Furthermore, after adjusting for sociodemographic factors, A-Levels and medical school admissions tests were all found to be statistically significant predictors of MRCS Part A success at the first attempt ($p < 0.05$).

Statistically significant correlations were found between all measures of medical school performance and MRCS Part A scores at the first attempt ($p < 0.001$).[34] EPM scores were found to correlate with Part A scores ($r = 0.57$), with performance deciles showing the strongest correlation ($r = 0.59$). EPM decile was found to be an independent predictor of MRCS Part A success after adjusting for sociodemographic factors, with the odds of passing Part A MRCS at first attempt increasing by 55% for every increase in EPM decile (odds ratio [OR] 1.55, 95% confidence interval [CI] 1.48 to 1.61).

The odds of passing Part A at the first attempt increased by 20% for every additional point awarded in the EPM for degree-level qualifications (OR 1.20, 95% CI 1.13 to 1.29). The odds of passing Part A on the first attempt also increased by 14% for every additional point awarded in the EPM for peer-reviewed publications (OR 1.14, 95% CI 1.01 to 1.28). SJT score was not found to independently predict Part A first attempt success after sociodemographic factors were taken into account ($P = 0.177$). This is perhaps unsurprising given that, unlike the other assessments, the FP-SJT is not a measure of knowledge or aptitude but a measure of personal attributes.

MRCS Part A score was found to correlate with MRCS Part B score ($r = 0.41$) and was an independent predictor of Part B success (OR 1.10 [95% CI 1.09 to 1.12]), as were the number of attempts taken to pass Part A (two vs one attempts OR 0.68 [95% CI 0.54 to 0.86], three vs one attempts OR 0.60 [0.43–0.84] and four or more vs one attempt OR 0.56 [95% CI 0.42 to 0.70]).[35] Part A score also correlated with FRCS Section 1 score ($r = 0.50$) and FRCS Section 2 score ($r = 0.34$). MRCS Part A score was found to independently predict FRCS Section 1 success (OR 1.14 [95% CI 1.09 to 1.89]) and FRCS Section 2 success (OR 1.06 [95% CI 1.01 to 1.11]) after adjusting for sociodemographic factors.[36]

Table 2 – Correlations between MRCS Part A Scores at the first attempt and medical school selection scores, FP selection scores, MRCS Part B and FRCS Section 1. A range of correlation coefficients is presented where the strength of correlation differs between cohorts sitting A-Levels and medical school admissions tests before and after 2010 when A* grades were introduced to A-Levels.

Source	Test	Pearson Correlation	P-Value	Number of Candidates
Ellis et al. **2021** ***Postgraduate Medical Journal***.[41]	**A-Level score**	0.17–0.22	<0.001	3235
	UCAT Total	0.25–0.26	<0.001	4515
	Verbal Reasoning	0.18–0.22	<0.001	
	Decision Making	0.14–0.20	<0.001	
	Quantitative Reasoning	0.26–0.27	<0.001	
	Abstract Reasoning	0.09–0.11	<0.001	
	BMAT Total	0.29–0.33	<0.001	3015
	Aptitude and Skills	0.25	<0.001	
	Scientific Knowledge and Applications	0.24–0.30	<0.001	
	GAMSAT Total	0.38	<0.001	395
	Reasoning in Humanities and Social Sciences	0.23–0.26	<0.001–0.007	
	Written Communication	0.05–0.12	0.051–0.635	
	Reasoning in Biological and Physical Sciences	0.38–0.41	<0.001	
Ellis et al. **2021 BMJ** ***Open***.[42]	**Educational Performance Measure**	0.57	<0.001	2585
	EPM Decile*	0.59	<0.001	
	EPM Degree Score	0.27	<0.001	
	EPM Publication Score	0.17	<0.001	
	Situational Judgement Test	0.23	<0.001	2585
Scrimgeour et al. **2018** ***The Surgeon***.[43]	**MRCS Part B**	0.41	<0.001	4310
Scrimgeour et al. **2019** ***BJS Open***.[44]	**FRCS** Section 1	0.50	<0.001	854
	FRCS Section 2	0.34	<0.001	797

* Spearman's Rho coefficient.

Discussion

Main findings

Statistically significant correlations were found between all examination scores included in the four studies reported in this paper, and MRCS Part A first attempt score. A-Levels, medical school admissions tests and performance at medical school all independently predicted MRCS Part A success. MRCS Part A scores were also found to be independently predictive of later success in MRCS Part B and FRCS Section 1. These findings agree with other studies that found a statistically significant correlation between MRCS Part A scores at the first attempt and scores at HST national selection for General and Vascular surgery ($r = 0.19$, $p < 0.001$).[38] Additionally, candidates passing Part A at the first attempt were found to be nearly twice as likely to remain in surgical careers (starting HST in surgical specialties) than candidates who failed at first attempt (OR 1.94 [95% CI (CI) 1.60 to 2.28] (P<0.001)).[1]

The combined data from these recent studies provide evidence of the predictive validity of the mandatory MRCS Part A for the first time. This is important to trainees, training programmes, and the Royal Colleges considering the international prestige and value currently placed on the MRCS within surgical training.

This programme of work joins a body of evidence from other UK postgraduate examinations. For example, the predictive validity of the MRCP has been assessed using each of its three components[32] and against assessment results taken before and after the MRCP including, but not limited to, A-Levels, UCAT, BMAT, the EPM and SJT, the PLAB and MRCGP examinations, with similar findings to those reported here.[15,16,39–41] Likewise, the predictive validity of the USMLE was assessed using the Medical College Admission Test, and performance in the American Board of Surgery qualifying and certifying examinations[12,13] reveals a similar relationship between examinations.

Whilst the current review has not included all examinations that a surgical candidate may attempt during their training, the combined data from the four studies represents most of the assessments commonly undertaken. Additionally, each of these studies involved statistically powerful, longitudinal analyses of large populations, thus enabling significant conclusions to be drawn from the data.

Limitations

Being a good doctor requires more than being successful in examinations and there are limitations to what examinations can effectively assess.[42] Examinations do fulfil a benchmarking function for each clinician's training and progression, often providing an objective assessment of knowledge that is used in parallel with the programmatic assessment of clinical competencies in the workplace. Associations between assessments taken throughout the training journey will reassure the public and key stakeholders and such associations could be used to partially justify the use, and role, of formal examinations throughout medical training. However, it is also important to accept the limitations of one-off high-stakes examinations in comparison to regular assessments of knowledge, skills and competencies within the clinical environment.[43]

An attempt was made to identify whether MRCS Part A performance could predict the risk of Fitness to Practice sanctions (FtP) by the GMC as an additional marker of future clinical performance. Poor performance in other postgraduate medical examinations including the MRCP, MRCGP, PLAB, American Board of Internal Medicine certification examinations and the USMLE is associated with an increased likelihood of sanctions and disciplinary action by medical regulators.[21,44–46] However, the reassuringly low number of GMC FtP sanctions within a large cohort of early-career surgeons in the UK (31 sanctions across 11,660 surgeons) prevented any meaningful statistical analyses of these variables.[47] A study of whether MRCS scores could predict 'on-the-job' performance in the form of Annual Review of Competence Progression (ARCP) outcomes revealed that only MRCS Part B was a statistically significant predictor of ARCP outcomes in HST, not MRCS Part A scores.[48] This was perhaps unsurprising as this was a comparison of two very different forms of assessment; a one-off, high-stakes, written examination of knowledge (MRCS Part A) and multiple measures of progression, mostly relying on clinical learning events supervised by senior clinicians, rather than a comparison of like-for-like. However, it is reassuring that MRCS Part B, a more similar assessment of knowledge and skills in the clinical environment, was predictive of ARCP outcomes. This will be discussed further in the second article of this series titled 'Establishing the Predictive Validity of the Intercollegiate Membership of the Royal Colleges of Surgeons Written Examination: MRCS Part B'. Furthermore, this similarity between assessment methods and examination performance may also explain the stronger correlation between EPM, a measure of performance in clinical examinations at Medical School, in comparison to weaker correlations seen between Part A and non-clinical medical school admissions tests. It is however difficult to assess the similarity of the assessments comprising the EPM score with MRCS Part A as currently UK medical schools have flexibility in how they examine their students. The introduction of the UK's Medical Licensing Examination (UKMLE) for students graduating in the academic year 2024–2025 onwards will provide a standardised comparison for future work examining the associations between different examinations in the medical and surgical training pathway.

There is little doubt amongst medical educationalists that candidates who perform well in early assessments continue to perform well in later assessments.[49] That these high-performing candidates continue to perform well at MRCS Part A is to be expected if there is predictive validity between assessments taken throughout a surgeon's educational career. All MRCS candidates are exceptional and have been high-achievers at school to be selected for medical school, and later CST. However, there will always be a spread of candidate scores at MRCS Part A. This provides trainers and training programmes with an opportunity to identify those at increased risk of failing future assessments in surgery such as MRCS Part B and FRCS Section 1, and provide additional training opportunities and support for future examinations.

Conclusion

Statistically significant correlations have been found between MRCS Part A and other assessments taken in the surgical training pathway. The combined data of the studies summarised in this review have provided compelling evidence of the predictive validity of this mandatory examination for the first time.

Funding

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

Contributorship

RE wrote the first draft of the manuscript. PB, JC, DS, AL and JH reviewed and edited the manuscript. All authors have approved the final copy.

Declaration of competing interest

None.

Acknowledgements

The authors would like to acknowledge Iain Targett at the Royal College of Surgeons of England, and Gregory Ayre from the Intercollegiate Committee for Basic Surgical Examinations for their support during this project.

Abbreviations

ARCP	Annual Review of Competence Progression
BMAT	Biomedical Admissions Test
CCT	Certificate of Completion of Training
CST	Core Surgical Training
EPM	Educational Performance Measure
FP	Foundation Training Programme
FP-SJT	Foundation Programme Situational Judgement Test (SJT)
FRCS	Fellowship of the Royal Colleges of Surgeons
FtP	Fitness to Practice
GAMSAT	Graduate Medical School Admissions Test
GMC	General Medical Council
ICBSE	Intercollegiate Committee for Basic Surgical Examinations
MCQ	Multiple Choice Questionnaire
MRCP	Membership of the Royal College of Physicians
MRCGP	Membership of the Royal College of General Practitioners
MRCS	Membership of the Royal Colleges of Surgeons
OSCE	Objective Structured Clinical Examination
PLAB	Professional and Linguistic Assessment Board Test
UCAT	University Clinical Aptitude Test (formerly the UKCAT)
UK	United Kingdom
USMLE	United States Medical Licensing Examination

REFERENCES

1. Ellis R, Cleland J, Scrimgeour D, Lee A, Brennan P. Does the MRCS exam fulfil its purpose in surgical professions? *Bull Roy Coll Surg Engl* 2021;**103**(7):344–50.
2. Nagy P. The three roles of assessment: gatekeeping, accountability, and instructional diagnosis. *Can J Educ Rev Can Léducation*. 2000;**25**(4):262.
3. Ellis R, Scrimgeour DSG, Brennan PA. The personal cost of postgraduate medical exams: are we asking too much of trainees?. Feb 2 [cited 2021 May 21];thebmjopinion. Available from: *Br Med J [Internet]* 2021. https://blogs.bmj.com/bmj/2021/02/02/the-personal-cost-of-postgraduate-medical-exams-are-we-asking-too-much-of-trainees/.
4. *Guide to the Intercollegiate Membership of the Royal College of surgeons examinations [Internet]. Intercollegiate Committee for Basic Surgical Examinations* [cited 2020 Sep 23]. Available from: https://www.intercollegiatemrcsexams.org.uk/-/media/files/imrcs/mrcs/mrcs-candidate-guidance/candidate-guide-to-mrcs-examination-january-2019-clean.pdf; 2018.
5. *The Standards for Educational and Psychological Testing*. Washington, DC, U.S: American Educational Research Association, American Psychological Association, National Council on Measurement in Education and Joint Committee on Standards for Educational and Psychological Testing; 2014.
6. Schuwirth LW. Assessing medical competence: finding the right answers. *Clin Teach* 2004;**1**(1):14–8.
7. Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ* 2004;**38**(9):1006–12.
8. *Intercollegiate Committee for Basic Surgical Examinations 2018/19 Annual Report [Internet]. Intercollegiate Committee for Basic Surgical Examinations* [cited 2020 Sep 23]. Available from: https://www.intercollegiatemrcsexams.org.uk/-/media/files/imrcs/policies-and-reports/icbse-annual-report-201819-final.docx; 2019.
9. Salkind N. *Encyclopedia of research design*, vol. 1. Thousand Oaks, CA: SAGE Publications, Inc.; 2010.
10. Hutchinson L, Aitken P, Hayes T. Are medical postgraduate certification processes valid? A systematic review of the published evidence. *Med Educ* 2002;**36**(1):73–91.
11. Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ* 2003;**37**(9):830–7.
12. de Virgilio C, Yaghoubian A, Kaji A, Collins JC, Deveney K, Dolich M, et al. Predicting performance on the American Board of Surgery qualifying and certifying examinations: a multi-institutional study. *Arch Surg Chic Ill 1960* 2010;**145**(9):852–6.
13. Gauer JL, Wolff JM, Jackson JB. Do MCAT scores predict USMLE scores? An analysis on 5 years of medical student data. *Med Educ Online* 2016;**21**:31795.
14. Tamblyn R, Abrahamowicz M, Dauphinee WD, Hanley JA, Norcini J, Girard N, et al. Association between Licensure examination scores and practice in primary care. *JAMA* 2002;**288**(23):3019–26.
15. Wakeford R, Denney M, Ludka-Stempien K, Dacre J, McManus IC. 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;**15**(1):1.
16. McManus IC, Wakeford R. PLAB and UK graduates' performance on MRCP(UK) and MRCGP examinations: data linkage study. *BMJ* 2014;**348**:g2621.
17. McCaskill QE, Kirk JJ, Barata DM, Wludyka PS, Zenni EA, Chiu TT. USMLE step 1 scores as a significant predictor of future board passage in pediatrics. *Ambul Pediatr Off J Ambul Pediatr Assoc* 2007;**7**(2):192–5.
18. Wenghofer E, Klass D, Abrahamowicz M, Dauphinee D, Jacques A, Smee S, et al. Doctor scores on national qualifying examinations predict quality of care in future practice. *Med Educ* 2009;**43**(12):1166–73.
19. Tamblyn R, Abrahamowicz M, Dauphinee D, Wenghofer E, Jacques A, Klass D, et al. Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *JAMA* 2007;**298**(9):993–1001.
20. Norcini JJ, Boulet JR, Opalek A, Dauphinee WD. The relationship between licensing examination performance and the outcomes of care by international medical school graduates. *Acad Med J Assoc Am Med Coll* 2014;**89**(8):1157–62.
21. Wakeford R, Ludka K, Woolf K, McManus IC. Fitness to practise sanctions in UK doctors are predicted by poor performance at MRCGP and MRCP(UK) assessments: data linkage study. *BMC Med* 2018;**16**(1):230.
22. Designing and maintaining postgraduate assessment programmes. *General Medical Council. [Internet]*. 2017 [cited 2020 Apr 28]. Available from: https://www.gmc-uk.org/education/standards-guidance-and-curricula/guidance/designing-and-maintaining-postgraduate-assessment-programmes.
23. Cleland J, Dowell J, McLachlan J, Nicholson S, Patterson F. Identifying best practice in the selection of medical students [cited 2020 Aug 2]; Available from: *Gen Med Counc [Internet]* 2012. http://www.gmc-uk.org/Identifying_best_practice_in_the_selection_of_medical_students.pdf_51119804.pdf.
24. *MSC Selection Alliance | Medical Schools Council [Internet]* [cited 2022 Dec 17]. Available from: https://www.medschools.ac.uk/our-work/selection/msc-selection-alliance.
25. *Educational Performance Measure (EPM) 2020 Framework [Internet]. UK Foundation Programme* [cited 2020 May 19]. Available from: https://foundationprogramme.nhs.uk/resources/; 2020.
26. *UKFP 2021 Applicants' Handbook [Internet]. UK Foundation Programme* [cited 2021 Jan 28]. Available from: https://foundationprogramme.nhs.uk; 2020.
27. *Good Medical Practice [Internet]. General Medical Council* [cited 2020 May 19]. Available from: www.gmc-uk.org/guidance; 2013.
28. *UK Foundation Programme 2020 Applicants Handbook. [Internet]* [cited 2020 May 19]. Available from: https://www.foundationprogramme.nhs.uk/wp-content/uploads/sites/2/2019/10/UKFP-2020-Applicants-Handbook-.pdf; 2020.
29. Joint Committee on Intercollegiate Examinations [Internet]. [cited 2020 Sep 29]. Available from: https://www.jcie.org.uk/.
30. Joint Committee on Surgical Training [Internet]. [cited 2020 Sep 29]. Certification Guidelines and Checklists. Available from: https://www.jcst.org/quality-assurance/certification-guidelines-and-checklists/.
31. *Intercollegiate Committee for Basic Surgical Examinations 2017/18 Annual Report [Internet]. Intercollegiate Committee for Basic Surgical Examinations* [cited 2020 Sep 30]. Available from: https://www.intercollegiatemrcsexams.org.uk/-/media/files/imrcs/policies-and-reports/icbseannual-report201718final.pdf; 2018.
32. 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;**10**(1):60.
33. Ellis R, Brennan P, Scrimgeour DS, Lee AJ, Cleland J. Performance at medical school selection correlates with success in Part A of the intercollegiate Membership of the Royal College of Surgeons (MRCS) examination. *Postgrad Med* 2021;**98**(1161). postgradmedj-2021-139748.
34. Ellis R, Scrimgeour DSG, Brennan PA, Lee AJ, Cleland J. Does performance at medical school predict success at the Intercollegiate Membership of the Royal College of Surgeons (MRCS) examination? A retrospective cohort study. *BMJ Open* 2021;**11**(8):e046615.
35. Scrimgeour DSG, Cleland J, Lee AJ, Brennan PA. 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;**16**(4):220–6.
36. Scrimgeour DSG, Cleland J, Lee AJ, Brennan PA. Prediction of success at UK Specialty Board Examinations using the mandatory postgraduate UK surgical examination. *BJS Open* 2019;**3**(6):865–71.
37. Hemphill JF. Interpreting the magnitudes of correlation coefficients. *Am Psychol* 2003;**58**(1):78–9.
38. Scrimgeour DSG, Cleland J, Lee AJ, Griffiths G, McKinley AJ, Marx C, et al. Impact of performance in a mandatory postgraduate surgical examination on selection into specialty training: performance in a postgraduate surgical examination and selection into specialty training. *BJS Open* 2017;**1**(3):67–74.
39. McManus I, Woolf K, Dacre J, Paice E, Dewberry C. The Academic Backbone: longitudinal continuities in educational achievement from secondary school and medical school to MRCP(UK) and the specialist register in UK medical students and doctors. *BMC Med* 2013;**11**(1):242.
40. McManus IC, Harborne AC, Horsfall HL, Joseph T, Smith DT, Marshall-Andon T, 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;**18**(1):136.
41. McManus IC, Woolf K, Harrison D, Tiffin PA, Paton LW, Cheung KYF, et al. Predictive validity of A-level grades and teacher-predicted grades in UK medical school applicants: a retrospective analysis of administrative data in a time of COVID-19. *BMJ Open* 2021;**11**(12):e047354.
42. Richards Jr JM, Taylor CW, Price PB, Jacobsen TL. An investigation of the criterion problem for one group of medical specialists. *J Appl Psychol* 1965;**49**(2):79–90.
43. Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach* 2011;**33**(6):478–85.
44. Tiffin PA, Paton LW, Mwandigha LM, McLachlan JC, Illing J. Predicting fitness to practise events in international medical graduates who registered as UK doctors via the Professional and Linguistic Assessments Board (PLAB) system: a national cohort study. *BMC Med* 2017;**15**(1):66.
45. Papadakis MA, Arnold GK, Blank LL, Holmboe ES, Lipner RS. Performance during internal medicine residency training and subsequent disciplinary action by state licensing boards. *Ann Intern Med* 2008;**148**(11):869–76.
46. Cuddy MM, Young A, Gelman A, Swanson DB, Johnson DA, Dillon GF, et al. Exploring the relationships between USMLE performance and disciplinary action in practice: a validity study of score inferences from a licensure examination. *Acad Med J Assoc Am Med Coll* 2017;**92**(12):1780–5.
47. Ellis R, Cleland J, Scrimgeour DSG, Lee AJ, Brennan PA. A cross-sectional study examining the association between

MRCS performance and surgeons receiving sanctions against their medical registration. *Surgeon* 2021;**20**(4):211–5.
48. Scrimgeour D, Brennan P, Griffiths G, Lee A, Smith F, Cleland J. 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;**100**(8):669–75.
49. McManus I, Dewberry C, Nicholson S, Dowell JS, Woolf K, Potts HW. Construct-level predictive validity of educational attainment and intellectual aptitude tests in medical student selection: meta-regression of six UK longitudinal studies. *BMC Med* 2013;**11**(1):243.