

BMJ Open Does the UKCAT predict performance on exit from medical school? A national cohort study

R K MacKenzie,¹ J A Cleland,¹ D Ayansina,² S Nicholson³

To cite: MacKenzie RK, Cleland JA, Ayansina D, *et al*. Does the UKCAT predict performance on exit from medical school? A national cohort study. *BMJ Open* 2016;**6**:e011313. doi:10.1136/bmjopen-2016-011313

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-011313>).

Received 28 January 2016
Revised 28 June 2016
Accepted 1 September 2016

ABSTRACT

Objectives: Most UK medical programmes use aptitude tests during student selection, but large-scale studies of predictive validity are rare. This study assesses the UK Clinical Aptitude Test (UKCAT: <http://www.ukcat.ac.uk>), and 4 of its subscales, along with individual and contextual socioeconomic background factors, as predictors of performance during, and on exit from, medical school.

Methods: This was an observational study of 6294 medical students from 30 UK medical programmes who took the UKCAT from 2006 to 2008, for whom selection data from the UK Foundation Programme (UKFPO), the next stage of UK medical education training, were available in 2013. We included candidate demographics, UKCAT (cognitive domains; total scores), UKFPO Educational Performance Measure (EPM) and national exit situational judgement test (SJT). Multilevel modelling was used to assess relationships between variables, adjusting for confounders.

Results: The UKCAT—as a total score and in terms of the subtest scores—has significant predictive validity for performance on the UKFPO EPM and SJT. UKFPO performance was also affected positively by female gender, maturity, white ethnicity and coming from a higher social class area at the time of application to medical school. An inverse pattern was seen for a contextual measure of school, with those attending fee-paying schools performing significantly more weakly on the EPM decile, the EPM total and the total UKFPO score, but not the SJT, than those attending other types of school.

Conclusions: This large-scale study, the first to link 2 national databases—UKCAT and UKFPO, has shown that UKCAT is a predictor of medical school outcome. The data provide modest supportive evidence for the UKCAT's role in student selection. The conflicting relationships of socioeconomic contextual measures (area and school) with outcome adds to wider debates about the limitations of these measures, and indicates the need for further research.

Strengths and limitations of this study

- This large-scale study is the first to link two national databases—the UK Clinical Aptitude Test (UKCAT) and the UK Foundation Programme (UKFPO).
- This is the first study examining the predictive validity of the UKCAT and the impact of participants' sociodemographic characteristics against national performance indicators on exit from medical school.
- Only medical students who were accepted into medical school and were in final year due to graduate in 2013 are included in this study.
- The analysis considers only simple predictor–outcome correlations.
- The reliability of the Educational Performance Measure (EPM) component of the UKFPO is unreported due to its diverse components, so we are therefore cautious in interpreting our findings.

ensure it measures what it claims to measure, and does so robustly and fairly, is an important issue in terms of ensuring the best applicants are selected into medicine.^{1–4} However, the holy grail of an objective yet comprehensive assessment for medical student selection remains elusive. Prior academic attainment, the traditional basis for medical school selection, is now recognised as insufficient given grade inflation⁵ and clear evidence of the impact of sociodemographic factors on educational attainment and academic progression.⁶ There is also a growing realisation that there is more to being a capable medical student or doctor than academic performance.^{7–9} These factors have led to the development of a range of innovative selection methods such as structured multiple mini interviews (MMIs),^{10–11} selection centres,¹² situational judgement tests (SJTs)^{13–14} and aptitude tests.¹⁵

In respect of one of these methods, aptitude tests, a recent systematic review of the effectiveness of selection methods in medical



CrossMark

For numbered affiliations see end of article.

Correspondence to

Dr RK MacKenzie; r.k.mackenzie@abdn.ac.uk

INTRODUCTION

Assessing the predictive validity and reliability of any medical school selection tool, to

education and training concludes that there is a lack of consensus among researchers on their usefulness due to mixed evidence supporting their predictive validity and fairness.² This conclusion reflects the current situation with the UK Clinical Aptitude Test (UKCAT), the aptitude test used in the selection process at the majority of UK medical schools.

Predictive validity studies have focused primarily on the relationship between UKCAT scores and performance on individual medical school examinations.^{16–20} These studies have reached varying conclusions: from suggesting that the test provides no significant prediction¹⁹ to claims of significant predictive ability.²⁰ These contradictory results may be due to relatively low numbers, and hence limited statistical power, and/or be associated with the outcome measures used, rather than the UKCAT itself (the psychometric properties of local assessment are rarely presented by authors, so there is no way of knowing if these are themselves robust). Moreover, the generalisable findings from small-scale studies are limited as medical schools in the UK differ in terms of their curricula, the length of programmes and the timing and nature of student assessments. A recent study aimed to overcome some of these limitations by examining the relationship between UKCAT scores and first year performance across 12 UK medical schools, found that UKCAT scores were significant predictors of outcome.²¹ These findings are reassuring but clearly indicated the need for even larger scale studies to see the national picture given the differences across schools, as well as follow-up into later undergraduate years and postgraduate training.

In terms of fairness, the small number of studies which focus on the fairness of the UKCAT also provide conflicting results. James *et al.*²² found that men, white students from higher socioeconomic status (SES) and independent or with grammar schooling gained higher marks in UKCAT (in the first cohort from 2006) and Tiffin *et al.*²³ found that male sex was more important in predicting the UKCAT scores. Turner and Nicholson²⁴ concluded that the UKCAT can facilitate the independent selection of appropriate candidates for interview (in that students with lower UKCAT scores are independently less likely to be selected for interview) but that it is not predictive of success at interview. Tiffin *et al.*²⁵ found that the UKCAT may lead to more equitable provision of offers to those applying to medical school from lower sociodemographic groups. However, Lambe *et al.*²⁶ found that students from non-independent schools are disadvantaged in terms of UKCAT results and preparation for applying to medical school. One possible reason for this is the differing ways in which the UKCAT is used within the medical school selection process.²⁷ However, notwithstanding this variation in usage, whether or not the UKCAT may disadvantage some candidate groups requires further investigation.

In short, notwithstanding substantial research efforts, which are considerable in comparison to other aptitude

tests and most other selection instruments, the predictive validity and fairness of the UKCAT remains unclear. Investigating these issues robustly is now possible because of three factors. First, given that the UKCAT was introduced in 2006, students for whom the UKCAT was part of the selection process are now graduating. Second, the introduction of a standardised national process for selection into the stage of medical training in the UK, which immediately follows medical school, the Foundation Programme (UKFPO) (<http://www.foundationprogramme.nhs.uk>), enables studies examining the predictive validity of the UKCAT in relation to performance on exit from medical school. Finally, there is enthusiasm within the UK to work across agencies to integrate large-scale databases which will enable a range of large-scale longitudinal studies addressing questions of predictive validity and fairness in medical selection and training in the UK.

The aim of this paper is twofold. First, to evaluate the predictive validity of the UKCAT in relation to the UKFPO selection process measures across a large number of diverse medical schools and programmes. Second, to examine the relationships between a broad range of sociodemographic factors and performance on the UKFPO selection process measures.

METHODS

Design

This was a quantitative study grounded in post-positivist research philosophy.^{28 29}

Study population

Our sample was the 2013 cohort of UK medical students from 30 UKCAT medical programmes who used the UKCAT in their 2006, 2007 and 2008 selection processes due to graduate in 2013. Students may have sat in the UKCAT in different years as they may have deferred entry, been on a 4-year, 5-year or 6-year course and may have repeated one or more years of medical school. This was the first cohort for whom both UKCAT and UKFPO indicators were available. All students applying to medical schools in the UK had to sit the UKCAT in these years (including international students) unless it was impractical for them to travel to their nearest test centre. The ways in which medical schools used the UKCAT can be found in Adam *et al.*²⁷

Data description

Working within a datasafe haven (to ensure adherence to the highest standards of security, governance and confidentiality when storing, handling and analysing identifiable data), routine data held by UKCAT and UKFPO were matched and linked.

The following demographic and pre-entry scores were collected: age on admission to medical school; gender; ethnicity; type of secondary school attended (ie, whether fee paying or non-fee paying); indicators of SES

or socioeconomic classification (SEC), including Index of Multiple Deprivation (IMD) which is based on post-code and where lower quintile scores represent more affluent areas, and the National Statistics Socioeconomic Classification (NSSEC) which is based on parental occupation; domicile (UK, European Union (EU) or overseas); and academic achievement prior to admission, out of a maximum of 360 points (the Universities and Colleges Admissions Service (UCAS) tariff, which is a UK-wide means of differentiating students based on grades from various qualifications, in use since 1992 for all British university course applications, see <http://web.ucas.com/>). Finally, the UKCAT scores were collected for entry into data analysis. The UKCAT cognitive scores are used in the medical selection process (a separate UKCAT non-cognitive domain was included for research purposes only⁹). The UKCAT total (cognitive) score is out of 3600 points. This total is comprised of four cognitive domains—abstract reasoning (AR), DA, quantitative reasoning and verbal reasoning (VR). Each has a maximum score of 900 points, summed to give the total UKCAT score out of 3600 points.

The outcome measures were UKFPO selection scores. All medical students studying in the UK who wish to enter the UKFPO obtain two indicators of performance: an Educational Performance Measure (EPM) and the score they achieve for a SJT.

The EPM is a decile ranking (within each school) of an individual student's academic performance across all years of medical school except the final year, plus additional points for extra degrees, publications, etc. The total EPM score is based on three components, with a combined score of up to 50 points:

- ▶ Medical school performance by decile (scored as 34–43 points);
- ▶ Additional degrees, including intercalation (up to 5 points);
- ▶ Publications (up to 2 points).

The UKFPO SJT is also scored out of 50 points. It^{30–33} is a type of aptitude test that assesses judgement required for solving problems in work-related situations. The SJT focuses on key non-academic criteria deemed important for junior doctors on the basis of a detailed job analysis (commitment to professionalism, coping with pressure, communication, patient focus and effective teamwork^{14 30}). It presents candidates with hypothetical and challenging situations that they might encounter at work, and may involve working with others as part of a team, interacting with others and dealing with workplace problems. In response to each situation, candidates are presented with several possible actions (in multiple choice format) that could be taken when dealing with the problem described. It is administered to all final year medical students in the UK as part of the foundation programme application process, is taken in examination conditions, and consists of 70 questions to be answered in 2 hours 20 min ([http://www.](http://www.foundationprogramme.nhs.uk/pages/medical-students/SJT-EPM)

[foundationprogramme.nhs.uk/pages/medical-students/SJT-EPM](http://www.foundationprogramme.nhs.uk/pages/medical-students/SJT-EPM)).

The EPM and SJT are summed to give the UKFPO score out of 100. No reliability score is available for the composite UKFPO.

Statistical analysis

All data were analysed using SPSS V.22.0. Pearson's or Spearman's rank correlation coefficients (dependent on distribution) were used to examine the linear relationship between each of SJT score and EPM and continuous factors such as UKCAT scores and preadmission academic scores and age. Pearson's correlation coefficients were computed to assess the relationship between continuous predictors and the outcome variables. Owing to the large sample size, small correlation coefficients were likely to reach statistical significance. Therefore, in terms of practical interpretation of the magnitude of a correlation coefficient, we defined, a priori, low/weak correlation as $r=0.00$ – 0.29 , moderate correlation as $r=0.30$ – 0.49 and strong correlation as $r\geq 0.50$. Two-sample t-tests, analysis of variance, Kruskal-Wallis or Mann-Whitney U tests were used to compare UKFPO indices across levels of categorical factors as appropriate.

Multilevel linear models were constructed to assess the relationship between the independent variables of interest: UKCAT cognitive total, UKCAT cognitive domains and demographic variables with each of the four outcomes (SJT, EPM decile, EPM total and UKFPO total). Fixed-effects models were fitted first and then random intercepts and slopes were introduced using maximum likelihood methods. Models were adjusted for identified confounders such as gender, age at admission, IMD quintiles, year the UKCAT examination was taken and whether or not the student attended a fee-paying school. Note that NSSEC and ethnicity were dropped from the models due to issues with non-convergence—this could occur for a variety of reasons and it is not always clear why, although one of the more common reasons is presence of tiny numbers in subgroups. What is clear though is that the models stabilised once these two variables were removed. Interactions between our primary variables and year of UKCAT examination were tested using Wald statistics and were dropped from the models if not significant at the 5% level. Nested models were compared using information criteria such as the log-likelihood statistics, Akaike's information criteria and Schwarz's Bayesian information criteria.

RESULTS

The matching exercise (between UKFPO data and UKCAT data) succeeded in linking 6804 (82%) of UKFPO records. Where candidates were graduating from a UKCAT Consortium school, the match rate was 83%; where candidates were graduating from a

Table 1 Descriptive statistics of the demographic variables of the sample

Total group n=6294	n (%)
Age in years at admission to medical school	
Median (IQR)	18 (18–19)
16–21	5070 (83.6)
Over 21	994 (16.4)
Total	6064
Missing	230
Gender	
Female	3522 (56.0)
Male	2772 (44.0)
Total	6294
Ethnic group	
Asian	1373 (21.8)
Black	153 (2.4)
Caucasian	4256 (67.7)
Mixed	206 (3.1)
Other	209 (3.3)
Total	6197
Missing	97
Ethnic group	
Non-Caucasian	1941 (31.3)
Caucasian	4256 (67.7)
Total	6197
Missing	97
Type of secondary school attended	
Comprehensive	1666 (28.8)
Further education college	267 (4.6)
Grammar	1137 (19.6)
Independent	1639 (28.3)
Non-UK	493 (8.5)
Sixth form college	572 (9.9)
Other	17 (0.3)
Total	5791
Missing	503
Type of secondary school attended	
Fee paying	1639 (28.3)
Non-fee paying	4152 (71.7)
Total	5791
Missing	503
IMD quintile	
1	2257 (36.4)
2	1432 (23.1)
3	1031 (16.7)
4	634 (10.2)
5	384 (6.2)
Non-UK	454 (7.3)
Total	6192
Missing	102
NSSEC score	
1	4733 (89.7)
2	140 (2.7)
3	282 (5.3)
4	47 (0.9)
5	74 (1.4)
Total	5276
Missing	1018
Domicile	
UK	5648 (92.0)

Continued

Table 1 Continued

Total group n=6294	n (%)
European Union	135 (2.2)
International	358 (5.8)
Total	6141
Missing	153
Receipt of UKCAT bursary	
Yes	116 (1.8)
No	6178 (98.2)
Total	6294

IMD, Index of Multiple Deprivation; NSSEC, National Statistics Socio-economic Classification; UKCAT, UK Clinical Aptitude Test.

non-Consortium school, the rate was 79%. However, UKCAT and UKFPO results were available for only 6294 UK medical students who started their final year of medical school in 2012, representing 80.6% of the total population of that year group of UK medical school graduates.

Table 1 shows the demographic profile of the sample. Most students were from the UK (n=5648, 92.0%). A majority of them (56%) were female and/or Caucasian (67.7%). Just under a third of students had attended a fee-paying school (28.3%—compared with the UK average of 7%). A majority of the graduating medical students were from higher SEC backgrounds (59.5% from IMD 1 or 2).

Table 2 provides an overview of candidate performance on the UKCAT cognitive tests, showing the domain and total scores. These were conducted in 2006, 2007 or 2008 depending on when the student entered medical school (see earlier).

In terms of outcome measures, as would be expected in a decile system such as the EPM, the percentage of graduating students within each decile per school were relatively constant (varying between 7.8 and 10.7). The median total EPM was 41.0 (out of a possible 34–50, IQR 38.0–44.0). SJT median was 40.9 (out of 0–50), IQR 38.7–42.9 and total UKFPO score median 81.9 (out of a possible 84–100), IQR 77.7–85.9. Approximately one-third (32.7%) of the sample had no additional EPM points, while 53.4% (n=3363) gained three or four points, which indicates they had either intercalated or entered medicine as an Honours graduate. Most (68.1%) did not gain any points for publications, while 23.3% gained one point, and gained 8.3% two points.

Univariate analysis

The relationship between demographic factors, UKCAT (cognitive) scores and UKFPO (EPM and SJT) scores are presented in **table 3**. EPM deciles are ranked in ascending order of magnitude (10 being the highest).

Correlations

There were highly significant and strong correlations between UKCAT cognitive total and the individual

Table 2 UKCAT cognitive domain scores and reliability indices

n=6276	Overall mean (SD)	2006 (n=2893)	*Reliability index 2006	2007 (n=2752)	*Reliability index 2007	2008 (n=631)	*Reliability index 2008
Abstract reasoning	625 (80.6)	631 (81.3)	0.74	614 (76.5)	0.65–0.69	648 (87.0)	0.65
Decision analysis	635 (93.4)	635 (95.0)	0.71	625 (88.9)	0.75–0.76	681 (90.5)	0.61
Verbal reasoning	625 (81.0)	627 (85.0)	0.86	619 (75.8)	0.79–0.83	649 (79.2)	0.80
Quantitative reasoning	647 (71.0)	630 (71.6)	0.58	656 (60.4)	0.53–0.59	682 (87.9)	0.58
Total UKCAT score	2533 (220.0)	2524 (224)	Not available	2514 (198)	0.87	2660 (243)	0.86

*As denoted by Cronbach's α .
UKCAT, UK Clinical Aptitude Test.

Table 3 Relationship between demographic variables and outcomes

Characteristics	Outcome variables							
	EPM decile	p Value	EPM total	p Value	SJT score	p Value	UKFPO total	p Value
Age at entry*; rank correlation (r)	0.034	0.008	0.126	<0.001	-0.054	<0.001	-0.057	<0.001
Gender; median (IQR)†								
Male	5 (3–8)	<0.001	38 (41–44)	<0.001	40.4 (38.3–42.3)	<0.001	81.1 (76.9–85.2)	<0.001
Female	6 (4–8)		38 (41–44)		41.3 (39.1–43.3)		82.4 (78.3–86.4)	
Ethnic group; median (IQR)†								
Caucasian	6 (4–9)	<0.001	41 (38–44)	<0.001	41.3 (39.2–43.3)	<0.001	82.8 (78.8–86.6)	<0.001
Non-Caucasian	5 (2–7)		40 (37–43)		39.9 (37.7–41.9)		79.9 (75.8–83.9)	
Type of secondary school attended; median (IQR)†								
Fee paying	6 (3–8)	<0.001	41 (38–44)	0.872	41.1 (39–43.2)	0.488	82.1 (78–86.1)	0.776
Non-fee paying	6 (3–8)		41 (38–44)		41 (38.9–43)		82.1 (77.7–86.1)	
IMD quintile; median (IQR)‡								
Most affluent—I	6 (4–8)	<0.001	41 (38–44)	<0.001	41.3 (38–44)	<0.001	82.4 (78.4–86.3)	<0.001
II	6 (3–8)		41 (38–44)		41.3 (38–44)		82.4 (78.3–86.5)	
III	6 (3–8)		41 (38–44)		41.0 (38–44)		81.8 (77.5–85.6)	
IV	6 (3–8)		41 (38–43)		40.6 (38–43)		81.0 (77.1–85.3)	
least affluent—V	5 (2–8)		40 (38–43)		40.2 (37–43)		80.3 (76.5–84.1)	
NSSEC; median (IQR)‡								
NSSEC categories								
1	6 (3–8)	0.072	41 (38–44)	0.006	41.1 (39.1–43.1)	<0.001	82.3 (78.4–86.2)	<0.001
2	6 (3–8)		40 (37–43)		41.1 (38.9–42.9)		80.8 (78.3–84.1)	
3	5 (3–8)		41 (38–44)		40.6 (38.3–42.8)		81.5 (77.5–84.9)	
4	5 (3–8)		40 (37–43)		39.9 (37.8–42.2)		80.8 (77.1–83.6)	
5	5 (3–7)		39 (37–42)		40.0 (37.8–42.0)		78.7 (76.5–83.6)	
Domicile categories; median (IQR)‡								
EU	6 (4–8)	<0.001	42 (40–45)	<0.001	40.2 (38.2–42.5)	<0.001	82.5 (79.2–85.9)	<0.001
International	4 (2–7)		39 (36–43)		39.3 (36.5–41.3)		78.8 (73.9–83.3)	
UK	6 (3–8)		41 (38–44)		41 (38.8–43.1)		82.0 (77.9–86)	

*Spearman's correlation coefficient.

†Mann-Whitney U test.

‡Kruskal-Wallis test.

EPM, Educational Performance Measure; EU, European Union; IMD, Index of Multiple Deprivation; NSSEC, National Statistics Socio-economic Classification; SJT, situational judgement test; UKFPO, UK Foundation Programme.

UKCAT domain scores ($r=0.64$ – 0.71 (not shown in tabular form)) as previously found in other studies.^{16 18 22} EPM decile and total score showed very little correlation with any of the UKCAT

cognitive tests or total. SJT showed a weak correlation with total UKCAT ($r=0.208$), also showing a weak correlation with the VR domain of the UKCAT ($r=0.216$; table 4).

Table 4 Correlation matrix of UKCAT scores and the four outcomes

	EPM decile	Total EPM	SJT score	UKFPO total
Abstract reasoning	0.096	0.148	0.111	0.164
Decision analysis	0.094	0.133	0.131	0.167
Quant. reasoning	0.079	0.061	0.100	0.102
Verbal reasoning	0.148	0.167	0.216	0.242
Total UKCAT score	0.155	0.193	0.208	0.253

EPM, Educational Performance Measure; SJT, situational judgement test; UKCAT, UK Clinical Aptitude Test; UKFPO, UK Foundation Programme.

Multivariate analysis

Contextually, UKCAT score is out of 3600 with a range of scores from 1630 to 3380 in our data (a spread of 1750 points). An increase of 333, 500, 333 and 167 in the total UKCAT cognitive score increases the average SJT, EPM decile, EPM total and UKFPO scores by one mark, respectively. In other words, every unit increase in UKCAT score results in a 0.006 of a unit increase in UKFPO—this represents a difference of 10.5 points on UKFPO between the highest and lowest UKCAT score, with the range of 1750. Small differences in UKCAT points translate into meaningful prediction in terms of the four outcomes of SJT, EPM decile, EPM total and UKFPO score.

Table 5 shows that UKCAT cognitive total score, all cognitive domain scores, except AR for the SJT (particularly VR), more affluent IMD quintile, female gender, increasing age and the year in which the UKCAT was attempted were significantly and positively associated with all four outcome measures (the EPM total, EPM decile, SJT and UKFPO total score). On the other hand, those who attended a fee-paying school did significantly *worse*, on average, than those who attended non-fee-paying schools. This weaker performance was particularly apparent on the total EPM and hence the UKFPO total score.

Each increase in IMD quintile (where higher quintiles indicate lower socioeconomic class) gives 0.357 fewer UKFPO points, an average difference of 1.43 UKFPO points between the highest and lowest IMD quintiles after correcting for other variables. For EPM decile, total EPM and SJT, each increase in IMD quintile gives 0.18, 0.23 and 0.13 fewer points, respectively.

Increasing age conferred an advantage, with an increase in UKFPO of one point for each additional 6 years of age. There was also a difference in which year the students sat their UKCAT, with those sitting in 2008 doing better than those in earlier years, most notably in UKFPO with a difference of 3.5 marks between 2007 and 2008.

DISCUSSION

This is the first study examining the predictive validity of the UKCAT and the impact of participants sociodemographic characteristics against national performance indicators on exit from medical school. Furthermore, it does so in a dataset representing 82% of the total population of the 2013 cohort of UK medical school final year students.

We have shown that the UKCAT—as a total score and in terms of the four cognitive subtests—has small but significant predictive validity of performance throughout medical school (indicated by the EPM). The UKCAT total score and three of the four cognitive subtests (not AR) also have small but significant predictive validity of performance on the SJT. Both of these outcome measures are used as the basis for selection to the next stage of medical training, the UKFPO. As the UKFPO selection process includes a medical school academic performance measure, this represents the first national comparison of the relationship between UKCAT scores and medical school outcomes, and hence is akin to the numerous studies investigating the predictive validity of medical school selection tests in countries with national licensure examinations, such as those exploring the relationship between performance on the US Medical College Admission Test and US Medical Licensing Examination (USMLE) steps.^{15 34 35}

It is crucial that any selection test is fair and adds value to the process. The SJT was designed to measure the attribute domains of commitment to professionalism, coping with pressure, effective communication, patient focus and working within teams as a first year doctor (<http://www.isfp.org.uk>), and so complement the academic ranking.^{30 36} Given this, should we have expected there to be a correlation between the UKCAT, an aptitude test and SJT, a test of the expression of job-specific personality traits in hypothetical situations? On the surface it would seem not, yet there is evidence that the ‘big five’ personality factors correlate with academic performance at medical school³³ and personal attributes such as motivation, resilience, perseverance and social and communication skills are associated with positive academic and work-related outcomes for young people (see Gutman and Schoon³⁷ for a recent review). The relationships between these different types of selection tool require further examination.

The patterns seen in the data align with the findings from earlier studies. We found that the UKCAT correlates with academic performance, which in this study is assessed by the EPM.^{17–21} Moreover, older students, female students and white students all perform better at medical school, with all effects being significant after taking educational attainment and UKCAT scores into account.^{16 17 36 38–43} The reasons for this differential performance of medical students remain unclear.

Our analysis adds that those living in more deprived areas on application (indicated by higher IMD quintile) also had significantly weaker performance on the SJT.

Table 5 Multilevel linear analysis

	SJT score		EPM decile		Total EPM		UKFPO score	
	Estimate (95% CI)	p Value	Estimate (95% CI)	p Value	Estimate (95% CI)	p Value	Estimate (95% CI)	p Value
Intercept	35.463 (33.682 to 37.243)	<0.001	1.549 (0.182 to 2.915)	0.026	31.686 (29.958 to 33.415)	<0.001	67 (64.262 to 69.738)	<0.001
Year of UKCAT examination								
2006	-0.446 (-0.914 to 0.023)	0.062	-0.588 (-0.948 to -0.229)	0.001	-0.087 (-0.542 to 0.368)	0.709	-0.478 (-1.197 to 0.241)	0.192
2007	-0.751 (-1.204 to -0.299)	0.001	-1.206 (-1.552 to -0.86)	<0.001	-3.004 (-3.443 to -2.565)	<0.001	-3.723 (-4.416 to -3.03)	<0.001
2008 and after	-		-		-		-	
Male gender	-0.964 (-1.157 to -0.771)	<0.001	-0.678 (-0.827 to -0.53)	<0.001	-0.724 (-0.911 to -0.536)	<0.001	-1.69 (-1.988 to -1.392)	<0.001
Attended fee-paying school	-0.24 (-0.45 to -0.029)	0.026	-0.46 (-0.622 to -0.298)	<0.001	-0.442 (-0.646 to -0.237)	<0.001	-0.672 (-0.997 to -0.347)	<0.001
IMD quintile	-0.132 (-0.211 to -0.054)	0.001	-0.178 (-0.239 to -0.118)	<0.001	-0.225 (-0.302 to -0.149)	<0.001	-0.36 (-0.481 to -0.238)	<0.001
Age at admission	-0.06 (-0.107 to -0.012)	0.014	0.044 (0.007 to 0.08)	0.019	0.265 (0.219 to 0.312)	<0.001	0.206 (0.132 to 0.279)	<0.001
Received UKCAT bursary	-0.061 (-0.773 to 0.651)	0.867	-0.126 (-0.674 to 0.423)	0.653	-0.404 (-1.095 to 0.286)	0.251	-0.469 (-1.569 to 0.631)	0.403
UKCAT score	0.003 (0.003 to 0.004)	<0.001	0.002 (0.002 to 0.002)	<0.001	0.003 (0.002 to 0.003)	<0.001	0.006 (0.005 to 0.006)	<0.001
Intercept	35.468 (33.679 to 37.257)	<0.001	1.444 (0.065 to 2.823)	0.04	31.592 (29.848 to 33.337)	<0.001	66.904 (64.15 to 69.658)	<0.001
Year of UKCAT examination								
2006	-0.47 (-0.94 to -0.001)	0.05	-0.608 (-0.97 to -0.246)	0.001	-0.111 (-0.57 to 0.348)	0.636	-0.52 (-1.242 to 0.201)	0.157
2007	-0.777 (-1.227 to -0.326)	0.001	-1.237 (-1.584 to -0.89)	<0.001	-3.033 (-3.473 to -2.594)	<0.001	-3.773 (-4.464 to -3.082)	<0.001
2008 and after	-		-		-		-	
Male gender	-0.999 (-1.195 to -0.803)	<0.001	-0.7 (-0.852 to -0.549)	<0.001	-0.738 (-0.929 to -0.547)	<0.001	-1.741 (-2.044 to -1.437)	<0.001
Attended fee-paying school	-0.227 (-0.437 to -0.017)	0.034	-0.453 (-0.615 to -0.291)	<0.001	-0.436 (-0.64 to -0.231)	<0.001	-0.652 (-0.976 to -0.328)	<0.001
IMD quintile	-0.118 (-0.196 to -0.039)	0.003	-0.173 (-0.234 to -0.113)	<0.001	-0.222 (-0.298 to -0.145)	<0.001	-0.342 (-0.463 to -0.22)	<0.001
Age at admission	-0.065 (-0.112 to -0.017)	0.007	0.041 (0.005 to 0.078)	0.027	0.263 (0.217 to 0.309)	<0.001	0.198 (0.125 to 0.271)	<0.001
Received UKCAT bursary	-0.018 (-0.727 to 0.692)	0.961	-0.109 (-0.656 to 0.439)	0.697	-0.39 (-1.081 to 0.301)	0.268	-0.413 (-1.51 to 0.684)	0.46
Abstract reasoning	0.001 (-0.0004 to 0.002)	0.156	0.002 (0.0005 to 0.003)	0.003	0.002 (0.0012 to 0.004)	<0.001	0.003 (0.0015 to 0.005)	0.001
Decision analysis	0.002 (0.001 to 0.003)	<0.001	0.001 (0 to 0.002)	0.037	0.002 (0 to 0.003)	0.007	0.004 (0.002 to 0.005)	<0.001
Quantitative reasoning	0.002 (0.001 to 0.004)	0.004	0.002 (0.001 to 0.003)	0.001	0.003 (0.001 to 0.004)	0.001	0.005 (0.003 to 0.007)	<0.001
Verbal reasoning	0.007 (0.006 to 0.008)	<0.001	0.003 (0.002 to 0.004)	<0.001	0.004 (0.003 to 0.005)	<0.001	0.011 (0.009 to 0.013)	<0.001

Total UKCAT score followed by cognitive score domains.

EPM, Educational Performance Measure; IMD, Index of Multiple Deprivation; SJT, situational judgement test; UKCAT, UK Clinical Aptitude Test.

An inverse pattern was seen for the type of secondary school attended. Those attending fee-paying schools performed significantly more weakly on the EPM decile, the EPM total and the total UKFPO score, but not the SJT, than those attending other types of school. (note that being in receipt of a UKCAT bursary was not significant but the numbers were so small (1.8%) that affected the statistical analysis). SJTs have been proposed as ‘fairer’ than academic and cognitively oriented assessment tools but close investigation of the literature highlights that, to date, studies of SJT fairness have looked mostly at ethnicity and gender.^{42–46} With the exception of Lievens *et al*,⁴⁷ this is the first study to include measures of social class in the analysis of SJT performance. Our findings do not indicate a clear pattern of performance or that the SJT under study is any less affected by socio-economic variables than more traditional assessments, but this clearly needs further investigation looking across cohorts and groups.

Those attending fee-paying schools performed significantly more weakly on the EPM decile, the EPM total and the total UKFPO score, but not the SJT, than those attending other types of school. This reflects the UKCAT-12 finding that those students coming from secondary schools where pupils tend to have higher preadmission grades do worse in first year at medical school.²¹ We are not sure why this is the case but one possibility, according to McManus *et al*,²¹ is higher-achieving secondary schools achieve their results in part by supporting students in ways which do not generalise or transfer to university where such support is not present. Interestingly, these data confirm that secondary school and IMD indicate or measure different factors, and add to the ongoing conversation about problems with all measures of SES.^{48–49} Moreover, given that UKFPO selection is obviously a competitive process, do lower scores on the various outcome measures mean that those medical students graduating from lower socioeconomic backgrounds may be less likely to get their first choice of programme? We need to follow-up this cohort, to investigate what this difference means in reality.

The strengths of this study are its large-scale, the integration of databases held by different agencies and its focus on the time of exit from medical school and selection to the next stage of postgraduate training in the UK. As with any study, there are limitations with this study. We have no way of knowing how the UKCAT was used to select the students in this sample, but it is likely that it was used in different ways by different medical schools.²⁷ The study is restricted to medical school entrants and, therefore, it cannot look more generally at how social and other factors relate to educational attainment and UKCAT performance in the entire set of medical school applicants. Only medical students who were accepted into medical school and were in final year in 2013 are included in this study but UKCAT routine data indicates that there was no difference in UKCAT scores in population of those applying to

medicine versus those who made it to final year, indicating that our sample is representative. The analysis also considers only simple predictor–outcome correlations. The outcome measures used—the UKFPO EPM and SJT are relatively new. However, we do have some indicators of their psychometric properties—predictive validity studies of the UKFPO SJT are just emerging, with the evidence indicating that higher SJT scores are associated with higher ratings of performance in foundation training.⁴³ While we do not know the reliability or predictive validity of the EPM and performance postmedical school, and, this necessitates caution when interpreting our findings, the wider literature indicates that performance in medical school exit examinations can predict performance in examinations during training.^{21 50 51} This requires follow-up, which will now be possible via the advent of the UK Medical Education Database (UKMED: <http://www.ukmed.ac.uk/>).

In conclusion, this large-scale study, the first to link two national databases (UKCAT and UKFPO), has shown that UKCAT is a predictor of medical school outcome. The data provide ongoing supportive evidence for the UKCAT’s role in student selection. The conflicting relationships of socioeconomic contextual measures (area and school) with outcome adds to wider debates about the limitations of these measures, and indicates the need for further research exploring how best to widen access to medicine.

Author affiliations

¹Institute of Education for Medical and Dental Sciences, School of Medicine and Dentistry, University of Aberdeen, Aberdeen, UK

²Department of Medical Statistics, College of Life Sciences and Medicine, University of Aberdeen, Aberdeen, UK

³Centre for Medical Education, Institute of Health Sciences Education, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Acknowledgements The authors thank the UKCAT Research Group for funding this independent evaluation and thank Rachel Greatrix of the UKCAT Consortium for her support throughout this project, and their feedback on the draft paper. The authors also thank the UKFPO for data provision and their support throughout. They also thank Professor Amanda Lee and Katie Wilde for their input into the application for funding, and ongoing support.

Contributors JAC and RKM wrote the funding bid. SN advised on the nature of the data. RKM managed the data and carried out the preliminary data analysis under the supervision of DA. DA advised on all the statistical analysis and carried out the multivariate analysis. JAC wrote the first draft of the Introduction and Methods sections of this paper. RKM and DA wrote the first draft of the Results section, RKM and SN wrote the first draft of the Discussion and Conclusions sections. JAC edited the drafts. All authors reviewed and agreed the final draft of the paper.

Funding United Kingdom Clinical Aptitude Test Consortium (grant number RG10984-10).

Competing interests This study addressed a research question posed and funded by UKCAT, of which SN was Chair. JAC is a member of the UKCAT Research Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Cleland JA, Dowell J, McLachlan J, *et al.* Identifying best practice in the selection of medical students. A project commissioned by the General Medical Council (GMC) and published in 2013 on the GMC website: http://www.gmc-uk.org/Identifying_best_practice_in_the_selection_of_medical_students.pdf. 51119804.pdf
- Patterson F, Knight A, Dowell J, *et al.* How effective are selection methods in medical education? A systematic review. *Med Educ* 2016;50:36–60.
- Prideaux D, Roberts C, Eva K, *et al.* Assessment for selection for the health care professions and specialty training: international consensus statement & recommendations. *Med Teach* 2011;33:215–23.
- Girotti JA, Park YS, Tekian A. Ensuring a fair and equitable selection of students to serve society's health care needs. *Med Educ* 2015;49:84–92.
- McManus IC, Powis DA, Wakeford R, *et al.* Intellectual aptitude tests and A-levels for selecting UK school-leaver entrants for medical school. *BMJ* 2005;331:555–9.
- Bowes L, Thomas L, Peck L, *et al.* *International research on the effectiveness of widening participation*. Higher Education Funding Council for England (HEFCE), 2013. <http://www.hefce.ac.uk/pubs/rereports/year/2013/wpeffectiveness/> (accessed 29 Sep 2015).
- Patterson F. Selection for medical education and training: research, theory and practice. In: Walsh K, ed. *Oxford textbook for medical education*. Oxford University Press, 2013:385–97; Chapter 33.
- MacKenzie RK, Dowell J, Ayansina D, *et al.* Do personal attributes assessed on medical school admission predict exit performance? A UK-wide longitudinal cohort study. *Adv Health Sci Educ* 2016; in press. doi:10.1007/s10459-016-9715-4
- Eva KW, Reiter HI, Rosenfeld J, *et al.* The ability of the multiple mini-interview to predict preclerkship performance in medical school. *Acad Med* 2004;79(10 Suppl):S40–2.
- Eva KW, Rosenfeld J, Reiter HI, *et al.* An admissions OSCE: the multiple mini-interview. *Med Educ* 2004;38:314–26.
- Eva KW, Reiter HI, Trinh K, *et al.* Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ* 2009;43:767–75.
- Gafni N, Moshinsky A, Eisenberg O, *et al.* Reliability estimates: behavioural stations and questionnaires in medical school admissions. *Med Educ* 2012;46:277–88.
- Patterson F, Carr V, Zibarras L, *et al.* New machine-marked tests for selection into core medical training: evidence from two validation studies. *Clin Med (Lond)* 2009;9:417–20.
- Patterson F, Zibarras L, Ashworth V. Situational judgement tests in medical education and training: research, theory and practice: AMEE Guide No. 100. *Med Teach* 2016;38:3–17.
- Callahan CA, Hojat M, Veloski J, *et al.* The predictive validity of three versions of the MCAT in relation to performance in medical school, residency, and licensing examinations: a longitudinal study of 36 classes of Jefferson Medical College. *Acad Med* 2010;85:980–7.
- Sartania N, McClure JD, Sweeting H, *et al.* Predictive power of UKCAT and other pre-admission measures for performance in a medical school in Glasgow: a cohort study. *BMC Med Educ* 2014;14:116.
- Husbands A, Mathieson A, Dowell J, *et al.* Predictive validity of the UK clinical aptitude test in the final years of medical school: a prospective cohort study. *BMC Med Educ* 2014;14:88.
- Yates J, James D. The value of the UKCAT in predicting pre-clinical performance: a prospective cohort study at Nottingham medical school. *BMC Med Educ* 2010;10:55.
- Lynch B, MacKenzie RK, Dowell J, *et al.* Does the UKCAT predict year 1 performance in medical school? *Med Educ* 2009;43:1203–9.
- Wright S, Bradley PM. Has the UK Clinical Aptitude Test improved medical student selection? *Med Educ* 2010;44:1069–76.
- McManus IC, Dewberry C, Nicholson S, *et al.* The UKCAT-12 study: educational attainment, aptitude test performance, demographic and socio economic contextual factors as predictors of first year outcome in a cross-sectional collaborative study of twelve UK medical schools. *BMC Med* 2013;11:244.
- James D, Yates J, Nicholson S. Comparison of A level and UKCAT performance in students applying to UK medical and dental schools in 2006: cohort study. *BMJ* 2010;340:c478.
- Tiffin PA, McLachlan JC, Webster L, *et al.* Comparison of the sensitivity of the UKCAT and A Levels to sociodemographic characteristics: a national study. *BMC Med Educ* 2014;14:7.
- Turner R, Nicholson S. Can the UKCAT select suitable candidates for interview? *Med Educ* 2011;45:1041–7.
- Tiffin PA, Dowell JS, McLachlan JC. Widening access to UK medical education for under-represented socioeconomic groups: modelling the impact of the UKCAT in the 2009 cohort. *BMJ* 2012;344:e1805.
- Lambe P, Waters C, Bristow D. The UK Clinical Aptitude Test: is it a fair test for selecting medical students? *Med Teach* 2012;34:e557–65.
- Adam J, Dowell J, Greatrix R. Use of UKCAT scores in student selection by UK medical schools, 2006–2010. *BMC Med Educ* 2011;11:98.
- Savin-Baden M, Howell Major C. *Qualitative research: the essential guide to theory and practice*. New York: Routledge, 2013.
- Cleland JA. Exploring versus measuring: considering the fundamental differences between qualitative and quantitative research. In: Cleland JA, Durning SJ. *Researching medical education*. Chichester: Wiley, 2015:3–14.
- Patterson F, Ashworth V. Situational judgment tests: the future of medical selection? *BMJ Careers* 2011. <http://careers.bmj.com/careers/advise/view-article.html?id=20005183> (accessed Sep 2015).
- Christian MS, Edwards BD, Bradley JC. Situational judgement tests: constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychol* 2010;63:83–117.
- Lievens F, Peeters H, Schollaert E. Situational judgment tests: a review of recent research. *Pers Rev* 2008;37:426–41.
- Lievens F. Adjusting medical school admission: assessing interpersonal skills using situational judgement tests. *Med Educ* 2013;47:182–9.
- Julian ER. Validity of the Medical College Admission Test for predicting medical school performance. *Acad Med* 2005;80:910–17.
- Donnon T, Paolucci EO, Violato C. The predictive validity of the MCAT for medical school performance and medical board licensing examinations: a meta-analysis of the published research. *Acad Med* 2007;82:100–6.
- Patterson F, Kerrin M, Edwards H, *et al.* Validation of the F1 Selection Tools, February 2015. Available via the UK Foundation Programme Website <http://www.foundationprogramme.nhs.uk/pages/home/about-the-foundation-programme> (accessed Sep 2015).
- Gutman LM, Schoon I. *The impact of non-cognitive skills on outcomes for young people: literature review*. London: Education Endowment Foundation, 2013. http://educationendowmentfoundation.org.uk/uploads/pdf/Non-cognitive_skills_literature_review.pdf (accessed Apr 2016).
- Yates J, James D. The UK clinical aptitude test and clinical course performance at Nottingham: a prospective cohort study. *BMC Med Educ* 2013;13:32.
- McManus IC, Woolf K, Dacer 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;11:242.
- Woolf K, Potts HWW, McManus IC. Ethnicity and academic performance in UK-trained doctors and medical students: systematic review and meta-analysis. *BMJ* 2011;342:d901.
- Esmail A, Roberts C. Academic performance of ethnic minority candidates and discrimination in the MRCPG examinations between 2010 and 2012: analysis of data. *BMJ* 2013;347:f5662.
- Whetzel D, McDaniel MA, Nguyen NT. Subgroup differences in situational judgment test performance: a meta-analysis. *Hum Perform* 2008;21:291–309.
- Wakeford R, Denney ML, Ludka-Stempien K, *et al.* Cross-comparison of MRCPG & 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;15:1–12.
- Motowidlo SJ, Tippins N. Further studies of the low fidelity simulation in the form of a situational inventory. *J Occup Organ Psychol* 1993;66:337–44.
- Schmidt F. The problem of group differences in ability test scores in employment selection. *J Vocat Behav* 1988;33:272–92.
- Chan D, Schmitt N. Video-based versus paper-and pencil method of assessment in situational judgement tests: subgroup differences in test performance and face validity perceptions. *J Appl Psychol* 1997;82:143–58.



47. Lievens F, Patterson F, Corstjens J, *et al.* Widening access in selection using situational judgment tests: evidence from UKCAT. *Med Educ* 2016;50:624–36. <http://onlinelibrary.wiley.com/doi/10.1111/medu.13060/abstract> (accessed May 2016).
48. Cleland JA, Patterson F, Dowell J, *et al.* How can greater consistency in selection between medical schools be encouraged? A mixed-methods programme of research that examines and develops the evidence base. A project commissioned by the Selecting for Excellence Group (SEEG) and published at <http://www.medschools.ac.uk/SiteCollectionDocuments/Selecting-for-Excellence-research-Professor-Jen-Cleland-et-al.pdf> (accessed Jul 2015).
49. Moore J, Mountford-Zimdars A, Wiggins J. Contextualised admissions: examining the evidence. A report to the SPA. 2013. http://www.spa.ac.uk/documents/ContextualData/SPA_ContextualisedAdmissions.July2013.pdf (accessed Sep 2015).
50. McDonald FS, Zegers SL, Kolar SE. Associations between United States Medical Licensing Examination (USMLE) and Internal Medicine In-Training Examination (IM-ITE) Scores. *J Gen Intern Med* 2008;23:1016–19.
51. Shellito JL, Osland JS, Helmer SD, *et al.* American Board of Surgery examinations: can we identify surgery residency applicants and residents who will pass the examinations on the first attempt? *Am J Surg* 2010;199:216–22.