BMJ Open

Does the UKCAT predict performance in medical and dental school? A systematic review

Rachel Greatrix ,[1] Sandra Nicholson,[2] Susan Anderson[3]

To cite: Greatrix R, Nicholson S, Anderson S. Does the UKCAT predict performance in medical and dental school? A systematic review. *BMJ Open* 2021;**11**:e040128. doi:10.1136/bmjopen-2020-040128

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-040128).

Received 05 May 2020
Revised 29 November 2020
Accepted 07 January 2021

[1]UCAT, School of Medicine, University of Nottingham, Nottingham, UK
[2]Centre for Medical Education, Queen Mary University of London, London, UK
[3]School of Life Sciences, University of Nottingham, Nottingham, UK

Correspondence to
Ms Rachel Greatrix;
Rachel.Greatrix@nottingham.ac.uk

ABSTRACT

Objectives For the first time, this systematic review provides a summary of the literature exploring the relationship between performance in the UK Clinical Aptitude Test (UKCAT) and assessments in undergraduate medical and dental training.

Design In accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis, relevant studies were identified through systematic literature searches. Electronic searches were carried out on EBSCO, EMBASE, Educational Resources Information Centre, SCOPUS, Web of Knowledge. Studies which included the predictive validity of selection criteria including some element of the UKCAT were considered.

Results 22 papers were identified for inclusion in the study. Four studies describe outcomes from dental programmes with limited results reported. 18 studies reported on relationships between the UKCAT and performance in undergraduate medical training. Of these, 15 studies reported relationships between the UKCAT cognitive tests and undergraduate medical assessments. Weak relationships (r=0.00–0.29) were observed in 14 of these studies; four studies reported some moderate relationships (r=0.30–0.49). The strongest relationships with performance in medical school were observed for the UKCAT total score and the verbal reasoning subtest. Relationships with knowledge-based assessments scores were higher than those for assessments of skills as the outcome. Relationships observed in small (single and double centre studies) were larger than those observed in multicentre studies.

Conclusion The results indicate that UKCAT scores predict performance in medical school assessments. The relationship is generally weak, although noticeably stronger for both the UKCAT total score and the verbal reasoning subtest. There is some evidence that UKCAT continues to predict performance throughout medical school. We recommend more optimal approaches to future studies. This assessment of existing evidence should assist medical/dental schools in their evaluation of selection processes.

Strengths and limitations of this study

- ► This study is the first to synthesise outcomes from articles reporting on the predictive validity of the UK Clinical Aptitude Test .
- ► Variability between study design, timelines and outcome markers creates challenges in synthesising data.
- ► The small number of studies relating to dental education prevented detailed analysis of those outcomes.

INTRODUCTION

In 2005, collaboration between 23 medical and eight dental school led to the development of the UK Clinical Aptitude Test (UKCAT). Since then a number of Universities have joined and left the UKCAT Consortium. In 2020, 30 medical schools used the test as part of their selection processes. In the same year, following international expansion, UKCAT changed its name to University Clinical Aptitude Test.

The UKCAT was designed to help select applicants to medical and dental programmes most suitable to undertake careers in medicine and dentistry. The test measures a range of cognitive skills identified by Consortium Universities as the skills required for success on medical and dental programmes. The Situational Judgement Test (SJT), introduced in 2013, seeks to assess a range of personal attributes required by successful clinicians. It is at the discretion of individual universities how they use the test in selection and its impact on selection has grown over time.[1]

A review in 2016 commissioned by the Medical Schools Council [The Medical Schools Council is the representative body for UK Medical Schools (https://www.medschools.ac.uk/]), investigated evidence underpinning selection in the UK.[2] The review focused separately on different selection criteria exploring issues relating to effectiveness, procedural issues, acceptability and cost-effectiveness. The main finding in relation to the use of aptitude test was one of conflicting evidence. Results varied between different tests, making generalised conclusions regarding their use difficult. Subgroup differences in performance were noted (such as gender, age and socioeconomic status), raising issues relating to fairness.

Given that the UKCAT is an established part of selection to medicine and dentistry in the UK, it is critical that we understand the ability of the test to predict performance on medicine and dental programmes and indeed professional performance beyond undergraduate training. While a number of predictive validity studies have been undertaken, no attempt has been made to date to synthesise these data. This systematic review will seek to provide a better understanding of the literature to assist end users of the test in making more informed decisions regarding selection processes.

Predictive validity studies are critical to help establish confidence in the use of selection tools to inform university selection processes. At the same time, candidates ought to be reassured of the legitimacy of measures which might otherwise be regarded as a further hurdle in selection.

Performance in the UKCAT is only one of many criteria used by universities in selection. Predicted and achieved academic measures, personal statements and references are (or have been) routinely used to identify applicants to be interviewed. The interview itself (whether structured, semistructured or Multiple Mini-Interview) will predominantly seek to identify those applicants with the correct personal qualities to pursue a successful healthcare career. It is, therefore, also of interest as to how each of these measures predict outcomes, the extent to which different criteria interact and overlap and, critically for UKCAT, its ability to predict outcomes over and above other criteria.

School leaving qualifications predict elements of performance in medical/dental schools and later postgraduate performance.[3] However, a combination of grade inflation and significant competition for places has led to a reduced ability to use these grades to discriminate between applicants.[4] UKCAT (and other admission tests), therefore, provide an opportunity to differentiate between high performing applicants with very similar academic records.

Longitudinal studies are difficult to undertake and by definition, cannot be undertaken until the desired outcome is available for the relevant cohort or group. In the early days of UKCAT's development, a number of studies were undertaken looking at the first cohorts of test takers in the early stages of their programmes e.g,[5 6] More recently, studies looking at later performance in medical school and the foundation year application stage have taken place.[7 8] While some studies were single or perhaps dual centre, others used data extracted from the UKCAT database [UKCAT has created a research database of candidate test scores and demographics. The database is held at the University of Dundee Health Informatics Centre] to look at much larger cohorts across many universities.[9 10]

Most studies focused on the ability of the test to predict performance in university assessments. These may include tests of knowledge or skills or mixed assessments combining both these elements. The foundation year application process utilises an educational performance measure, which has also provided a useful outcome for analysis. The studies use a variety of outcome markers, at different stages of education and training. Researchers had access to different cohorts and different demographic variables.

The primary aim of this review is to evaluate existing evidence regarding the predictive validity of the UKCAT. Secondary outcomes envisaged include identification of more optimal approaches to future studies, identifying how cohorts might be best identified and outcome markers defined alongside appropriate methodologies. Information from the review may provide additional information to medical/dental schools in evaluating selection processes.

METHODS

The review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist.[11]

Selection of studies

The initial search and abstract screen took place in early 2018 and included searches on EBSCO, EMBASE, Educational Resources Information Centre, SCOPUS, Web of Knowledge using the following search terms:

((UKCAT)OR(UK CAT)OR(United Kingdom Clinical Aptitude Test)OR(UK Clinical Aptitude Test) AND((valid*)OR (predict*)OR(criteri*))

The search was restricted to studies after 2006 (the first year of delivery of UKCAT).

All identified titles/abstracts were collated and reviewed for relevance in relation to inclusion/exclusion criteria. Full-text papers were accessed for relevant studies and a further decision made regarding inclusion.

Studies were included if they contain predictive validity studies around selection to medical and dental education which included the UKCAT; where all or some of the analysis focused on the predictive validity of the UKCAT; where the target population was UKCAT test takers subsequently enrolled on medicine and dental programmes in the UK. Inclusion/exclusion criteria are summarised in table 1.

Specific searches using the search terms of the e-journal versions of Medical Education (2006–2017), Medical Teacher (2006–2017), Advances in Health Sciences Education (2006–2017) and BioMed Central (BMC) Medical Education (2006–2017) took place as well as a review of the available published abstracts of conference proceedings of the Annual Scientific Meeting of Association for the Study of Medical Education (ASME) and the Association for Medical Education Europe (AMEE) annual conferences.

Universities may have undertaken local analysis which has not been published. A request to access such reports/

Table 1 Inclusion/exclusion criteria

Inclusion criteria	Exclusion criteria
Empirical data	Not empirical data
Study population includes UKCAT test takers	Study population does not include UKCAT test takers
Study includes predictive validity of selection criteria including some element of the UKCAT	Study not focused on predictive validity of selection criteria
	Selection criteria do not include any element of the UKCAT

UKCAT, UK Clinical Aptitude Test.

analysis was made to UKCAT Consortium Universities. This did not result in additional data sources being identified.

Search results

Initial searches were undertaken by RG with outcomes verified independently by SA. The initial search took place in March 2018 with a further search in July 2018. The outcomes from the search of databases is included in figure 1.

Outcomes of a review of the grey literature (AMEE abstracts 2007–2017, Ottawa abstracts 2014, 2016, ASME abstracts 2009–2017, International Network for Researchers in Selection into Healthcare (INRESH) programmes, UKCAT Consortium Agendas) is included in figure 2.

Data extraction

Data were extracted by RG into a table and included article title, year of publication, sample size, number of universities included in the study, programme (medicine or dentistry) and year of admission.

Correlations between predictor and outcome variables were extracted and recorded. Effect sizes and 95% CIs were calculated for each relationship identified using the Practical Meta-Analysis Effect Size Calculator.[12] Where appropriate, effect sizes and 95% CIs were aggregated by UKCAT subtest (and total score) and programme year.

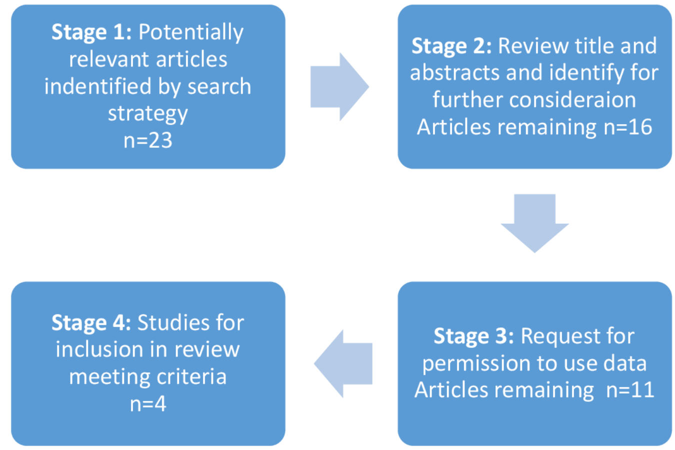

Figure 2 Flow diagram of outcomes from review of grey literature.

This allowed forest plots to be generated to illustrate relationships between each of the UKCAT subtests (and total score) and assessment measures over different years of study.

Similar to other studies,[13] when commenting on outcomes, correlation coefficients in the range 0.00–0.29 were defined as weak, r=0.30–0.49 as moderate. No strong correlations (r=0.5+) were observed in any of these studies.

Patient and public involvement

Patients were not involved in this study.

RESULTS

Twenty-two papers were identified for inclusion in the study; 18 published articles and four articles sourced from the grey literature.

Data available to researchers

Predictor variables

Predictor variables used were UKCAT total and subtest scores. The UKCAT comprises four cognitive subtests (Verbal Reasoning, Quantitative Reasoning, Abstract Reasoning and Decision Analysis) and (since 2013) a test of SJT. The cognitive tests are each reported using scaled

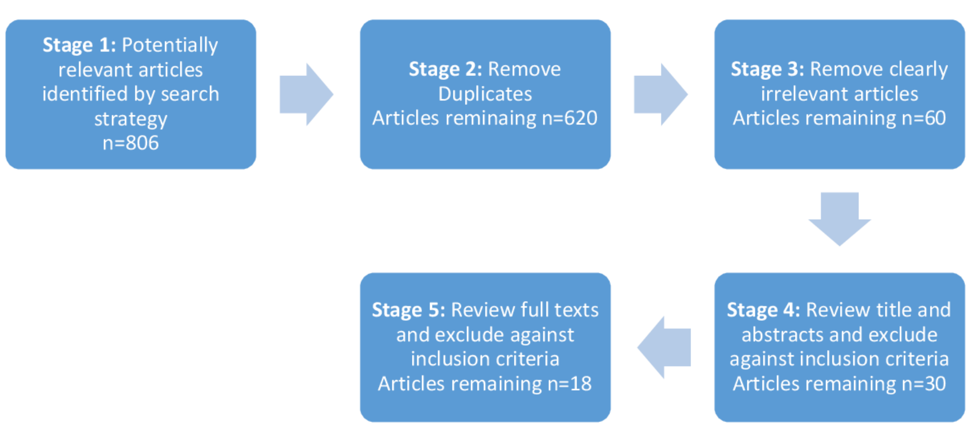

Figure 1 Flow diagram of paper selection process using databases.

scores (in the range 300–900) creating a UKCAT total score in the range of 1200–3600. Total score has been most commonly used in selection. The SJT is reported to candidates and universities (for use in selection) as a band (bands 1–4 with band 1 being the highest).

Many of the studies looked at a broad range of other selection parameters. Academic scores (some locally derived[14 15]), personal statement/reference scores, interview scores were often included in studies to build up an authentic model of selection. Many studies considered a range of demographic factors alongside these selection criteria. Commonly used factors were gender, age, social class, ethnicity and school type.

Other predictor variables included a test of non-cognitive traits which contained content similar to that trialled within the UKCAT in earlier years.[14 15]

Outcome measures

Outcome variables extracted from these studies were year of assessment, assessment outcome measure (eg, objective structured clinical examination (OSCE), Final Exam) and a coded proxy of the exam type (skills, knowledge or mixed assessments). The coded proxy drew on information presented in each paper as to the nature of assessment outcomes used. End of year and final examinations often include a combination of knowledge and skills assessments. The details of the weighting of the two elements were rarely reported. It is likely, however, that the knowledge element had greater weighting in 'mixed' assessment outcomes.

Outcome measures varied significantly between studies. Many single/dual centre studies utilised tutor ratings and summative assessments alongside measures of non-cognitive traits such as objective structured long examination records and OSCEs. Some studies stratified student performance in different ways, using measures such as grade boundaries,[16] graduation with honours[14 15] and fitness to practice penalty points.[14] The studies which worked on UKCAT's own national dataset (drawing on progression data obtained from medical and dental schools) used aggregated end of year knowledge, skills and total marks. For these studies, researchers did not have access to information to allow more detailed interpretation of these assessments. One study used bespoke supervisor ratings against which to measure the performance of the UKCAT SJT.[17]

Dental outcomes

Four papers using dental data were identified for inclusion, reporting only five statistical significant results ($p<0.05$) which are summarised in table 2.

Foley at al drew on admissions data from 2010 to 2014 but did not report how changes to mean average UKCAT scores across those years (eg, mean average 2010=2489, 2013=2642) were accounted for. Researchers reported a significant but small relationship between UKCAT percentile and assessments score ($r=0.118$, $p=0.001$).

Lala *et al*[18] only found significant correlations between decision analysis and assessment outcomes.

The relationships observed by Lambe *et al*[19] were moderate (between total UKCAT score and year 1 outcomes), although no significant relationships were found between subtest scores and assessment outcomes. The outcome measure here was described in the paper as 'academic knowledge of dental practice' but without further detailed information regarding the assessment it was not possible to explore this relationship further.

McAndrew *et al*[16] found no significant correlations between UKCAT and examination performance in year 1 at both Newcastle and Cardiff Dental schools. The study did, however, identify associations between UKCAT score and poor performance (determined by grade boundaries).

It was difficult to draw conclusions from the findings reported for dentistry given the small number of both studies and significant results. From this point, analysis will focus on studies involving medical students only.

Medicine outcomes

Of the remaining 18 studies, two papers reported exclusively on how the UKCAT SJT predicted performance in medical school.[17 20] Sixteen studies reported correlations between UKCAT (cognitive subtests) and knowledge, skills and mixed assessments (coded proxy of exam type).

Table 2 Dental schools: UKCAT predictive validity coefficients (R) with assessment outcome measures

			UKCAT predictor (r)					Outcome predictor
Study source	**N**	**University (year of admission)**	**VR**	**QR**	**AR**	**DA**	**Total Score**	**year K, S, M**
Lambe *et al*[19]	44	Peninsula (2014)	ns	ns	ns	ns	0.32	Yr 1 K
			ns	ns	ns	ns	0.38	Yr 1 K
Foley *et al*[28]	71	Aberdeen 2010, 2011, 2012, 2013, 2014					0.077	All M
Lala *et al*[18]	135	Sheffield (2008, 2009)	ns	ns	ns	0.203	ns	Yr 1 M
			ns	ns	ns	0.179	ns	Yr 1 M
McAndrew *et al*[16]	164	Cardiff, Newcastle					ns	Yr 1 M

$P<0.05$; ns indicates where relationships were explored but results were not significant; blank cells indicate that a relationship was not explored.
UKCAT, UK Clinical Aptitude Test.

Of these, 15 studies reported Pearson correlations or (within regression analyses) standardised regression coefficients (beta). These outcomes can be interpreted similarly to a correlation coefficient r.[21]

These 15 studies are presented in table 3 which records statistical significant ($p<0.05$) correlations between UKCAT (cognitive subtests and total score) and medical school assessment outcomes.

Sample sizes ranged from 44 to 6294 with an approximate total of 23 000 candidates/applicants included in these studies. Twelve of the studies were single centre. Five studies included a larger number of universities, where authors drew on national datasets (UKCAT and UK Medical Education Database (UKMED) [The UK Medical Education Database (UKMED) is a partnership between data providers from across education and health sectors supporting the creation of a database to analyse issues relating to selection, medical education and training and impact on career pathways. https://www.ukmed.ac.uk/]). Twelve of the studies used data from the first years of UKCAT delivery (entry to medical school in 2007 and 2008) with the most recent study drawing on data for 2014 entry.

Most studies looked at more than one programme year. Twelve of the studies used year one outcome data with seven studies using year 3, six studies using year 3, eight studies using year 4 and nine studies using year 5. The highest number of relationships were observed in years 5 and 1. The number of significant correlations identified were lowest for Abstract Reasoning and highest for UKCAT total score. The number of correlations observed for knowledge-based assessments were higher than for skills or mixed assessments.

The study reporting no significant relationships[5] investigated relationships between UKCAT and year 1 medicine outcomes at Aberdeen for the 2007 entrants. This finding was also reported in a later study[22] looking at the same cohort in which only one weak relationship between UKCAT Total Score and assessment outcomes was found. Interestingly, further analysis of this cohort in later years of medical school showed a moderate relationship between UKCAT total score and some year four and five assessment outcomes.[7]

Relationships between UKCAT and medical school assessments

Effect sizes and 95% CIs were calculated for each relationship identified in table 3 using the Practical Meta-Analysis Effect Size Calculator.[12] Effect sizes and 95% CIs were then aggregated by UKCAT subtest (and total score) and programme year. This allowed forest plots to be generated to illustrate the relationships between each of the UKCAT subtests (and total score) and assessment measures over different years of study. These are illustrated in the figures below.

Figure 3 illustrates the aggregated relationships between each UKCAT subtest (and total score) with all assessments included in the studies for each programme year.

The strongest relationships with all assessment outcomes were observed for UKCAT total score and verbal reasoning although all relationships were weak. There was a very small upwards trend in relationships over the 5 years, with slightly larger trends observed for UKCAT Total Score and Verbal Reasoning.

Figure 4 reports relationships between UKCAT subtests and assessments of knowledge. The strongest relationships with knowledge assessment outcomes were observed for UKCAT total score and verbal reasoning, although all relationships were weak. Relationships remain fairly constant over the 5 years. The relationships with knowledge assessments were generally higher than those observed with all assessments.

Figure 5 reports relationships between UKCAT subtests and skills assessments. The strongest relationships with skills assessment outcomes were observed for UKCAT total score and verbal reasoning, although all relationships were weak. Relationships with quantitative reasoning, abstract reasoning and decision analysis were low. There was a slight upwards trend in relationships over the 5 years observed in for UKCAT Total Score, verbal reasoning and abstract reasoning. Relationships are lower than those observed for knowledge-based assessments although the upwards trend was more noticeable.

Figure 6 illustrates the relationships between UKCAT subtests and mixed assessments (those involving assessments of both knowledge and skills). The strongest relationships with mixed assessments were observed for UKCAT total score and verbal reasoning, although all relationships were weak.

Figures 3–6 suggest that the strongest relationships with assessments outcomes were observed for verbal reasoning and total score although all relationships were weak. Relationships with skills assessments were weaker than for other assessment outcomes. There was some evidence of an upwards trend in relationships over programme years but again this was small and varied across subtests.

Single/double centre studies versus multicentre studies

Each study was identified as either small (single or dual centre) or large (multi-centre studies). Effect sizes were then aggregated (for UKCAT Total Score only) by programme year and outcome measure (knowledge, skills). This allowed forest plots to be generated to illustrate differences between relationships for studies of different sizes. These are illustrated in figure 7.

The relationships observed for both knowledge and skills assessments were stronger for studies involving a smaller number of Universities. Differences between outcomes for small and large scale studies were more noticeable for skills rather than knowledge based assessments. There was an increasing relationships overtime with skills outcomes in larger scale studies.

Table 3 Characteristics of the studies relating UKCAT with assessment outcome measures— knowledge (K), skills (S), mixed (M)

Study source	N	University (year of admission)	UKCAT predictor (r) VR	QR	AR	DA	Total Score	Outcome predictor year K, S, M
Adam *et al*[14]	146	Hull York Medical School (2007)	ns	ns	ns	ns	0.181	Yr 4 K
			ns	ns	0.231	ns	0.175	Yr 5 K
			0.244	ns	0.25	ns	0.204	Yr 5 s
Adam *et al*[15]	146	Hull York Medical School (2007)	0.363	0.233	0.234	0.181, 0.18	0.39, 0.214	Yr 1 K
			ns	ns	ns	0.197	ns	Yr 1 s
			ns	ns	0.255	0.209	0.277	Yr 1 M
			0.253	0.267	0.202	0.258, 0.204	0.212	Yr 2 K
			0.331	0.275	ns	0.281	0.377	Yr 2 s
			0.241	0.2	ns	0.291	0.323	Yr 2 M
Hanlon *et al*[22]	341	Aberdeen (2007, 2009)	ns	ns	ns	ns	0.167	Yr 1 K
Husbands and Dowell[26]	147	Dundee (2009)					0.25	Yr 1 K
							0.18	Yr 1 s
Husbands *et al*[7]	341	Aberdeen, Dundee (2007)					0.34, 0.24	Yr 4 K
							0.36	Yr 4 s
							0.29	Yr 5 s
Lynch *et al*[5]	341	Aberdeen, Dundee (2007)	ns	ns	ns	ns	ns	Yr 1 K
			ns	ns	ns	ns	ns	Yr 1 M
MacKenzie *et al*[8]	6294	All (2007, 2008, 2009)	0.216	0.1	0.111	0.131	0.208	Yr 5 s
			0.242, 0.167, 0.148	0.102, 0.079, 0.061	0.164, 0.148, 0.096	0.167, 0.133, 0.094	0.253, 0.196, 0.155	Yr 5 M
McManus *et al*[9]	4811	12 universities (2007, 2008, 2009)	0.177	0.079	0.052	0.077	0.16	Yr 1 K
			ns	0.044	0.053	0.056	0.075	Yr 1 s
			0.115	0.076	0.08	0.09	0.148	Yr 1 M
Sartania *et al*[25]	189	Glasgow (2007)	0.174	0.197	ns	0.172	0.252, 0.149	Yr 1 K
			0.145	0.155	ns	ns	0.187	Yr 5 K
			0.213, 0.201	0.219, 0.216	ns	0.174	0.216, 0.251	Yr 5 M

Continued

Table 3 Continued

Study source	N	University (year of admission)	UKCAT predictor (r) VR	QR	AR	DA	Total Score	Outcome predictor year K, S, M
Tiffin *et al*[23]	6425	18 universities (2007, 2008)	0.153	0.072	0.098	0.086	0.172	Yr 1 K
			0.065	0.039	0.06	0.065	0.1	Yr 1 s
			0.163	0.081	0.065	0.089	0.167	Yr 2 K
			0.072	0.021	0.09	0.067	0.113	Yr 2 s
			0.209	0.116	0.064	0.11	0.207	Yr 3 K
			0.111	0.045	0.052	0.072	0.12	Yr 3 s
			0.194	0.099	0.096	0.079	0.196	Yr 4 K
			0.13	0.062	0.097	0.075	0.154	Yr 4 s
			0.188	0.108	0.11	0.11	0.217	Yr 5 K
			0.131	0.072	0.096	0.082	0.161	Yr 5 s
Yates and James[29]	204	Nottingham (2007)	0.319, 0.189	0.24, 0.152	ns	ns	0.232, 0.211	Yr 2 K
			ns	ns	ns	−0.155	ns	Yr 2 s
Yates *et al*[30]	193	Nottingham (2007)	0.215	0.173	ns	ns	0.192	Yr 3 K
			0.188	ns	ns	ns	ns	Yr 3 s
			0.237	ns	ns	ns	0.173	Yr 3 M
			0.266	ns	ns	ns	0.176	Yr 4 K
			0.224	ns	ns	ns	0.259	Yr 4 s
			0.275	ns	ns	ns	0.242	Yr 4 M
			0.255	0.203	ns	ns	0.205	Yr 5 K
			ns	ns	ns	ns	ns	Yr 5 s
			0.237	0.183	ns	ns	0.193	Yr 5 M
Mwandigha *et al*[31]	2107	18 universities (2008)					0.11	Yr 1 K
							0.07	Yr 1 s
							0.11	Yr 2 K
							0.06	Yr 2 s
							0.15	Yr 3 K
							0.06	Yr 3 s
							0.11	Yr 4 K
							0.07	Yr 4 s
							0.16	Yr 5 K
							0.11	Yr 5 s

Continued

Table 3 Continued

Study source	N	University (year of admission)	UKCAT predictor (r) VR	QR	AR	DA	Total Score	Outcome predictor year K, S, M
Srikathirkamanathan and MacManus[32]	183	Southampton (2007, 2008)	0.278, 0.176	0.26, 0.146	ns	0.162	0.337, 0.25	Yr 5 K
			0.204, 0.157	ns	ns	ns	0.241	Yr 5 s
			0.285	0.195	ns	0.155	0.329	Yr 5 M
Tiffin and Paton[24]	1400	8 universities (2013)	0.14	0.16	ns	0.17	0.17	Yr 1 K
			ns	ns	ns	0.15	ns	Yr 1 s

P<0.05; ns indicates where relationships were explored but results were not significant; blank cells indicate that a relationship was not explored.
UKCAT, UK Clinical Aptitude Test.

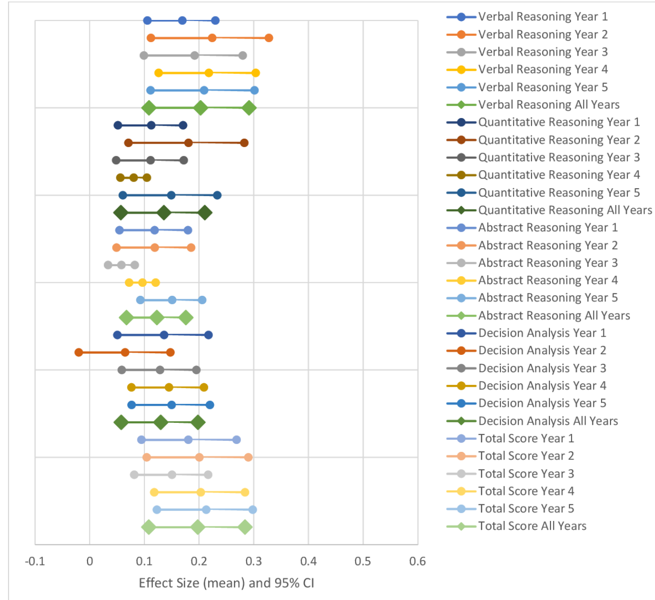

Figure 3 Relationships between UKCAT cognitive Subtests and all assessments. UKCAT, UK Clinical Aptitude Test.

Additional analysis

Many studies included alternative and/or additional analyses in addition to results identified above. These are discussed below.

Regression analyses

Table 4 summarises outcomes from regression analyses reported in some of the studies. Regression analyses

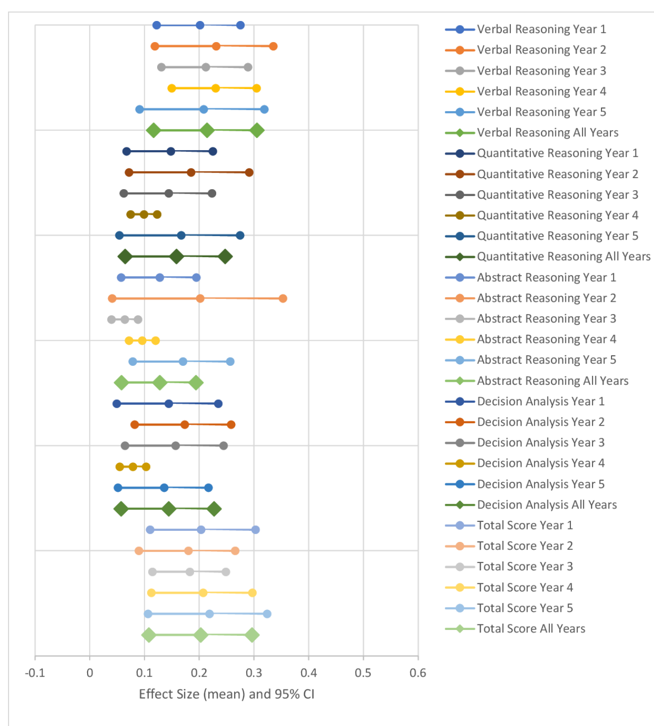

Figure 4 Relationships between UKCAT cognitive Subtests and assessments of knowledge. UKCAT, UK Clinical Aptitude Test.

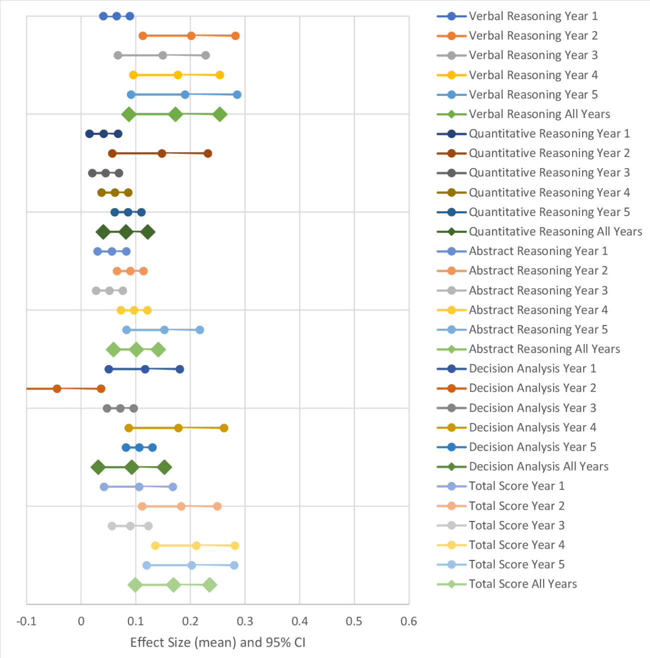

Figure 5 Relationships between UKCAT cognitive Subtests and assessments of skills. UKCAT, UK Clinical Aptitude Test.

examine whether a set of predictor variables predict the outcome variable. It is then possible to determine which variables are significant predictors of the outcome variable and the extent to which they impact the outcome variable. These analyses took a number of forms and, due to the diversity in both input and output variables, it was difficult to compare results. For example Adam *et al*[14] report a series of linear regressions to identify the best predictors of performance in years 4 and 5, whereas MacKenzie *et al*[13] use multilevel models to report predictors of performance in United Kingdom Foundation Programme (UKFPO) applications.

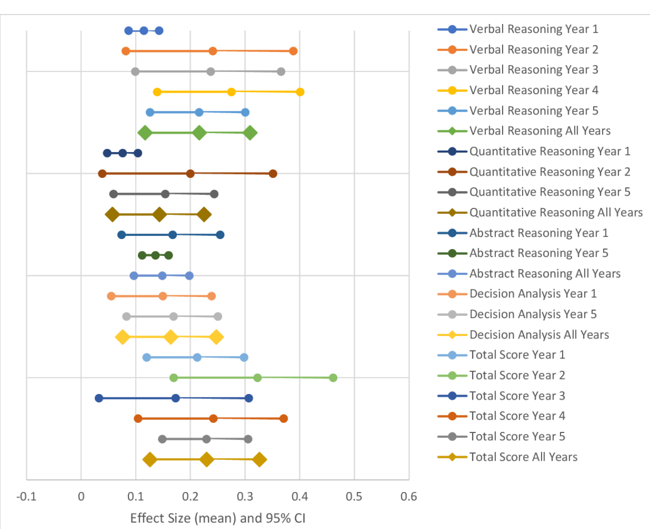

Figure 6 Relationships between UKCAT cognitive Subtests and mixed assessments (combining both knowledge and skills). UKCAT, UK Clinical Aptitude Test.

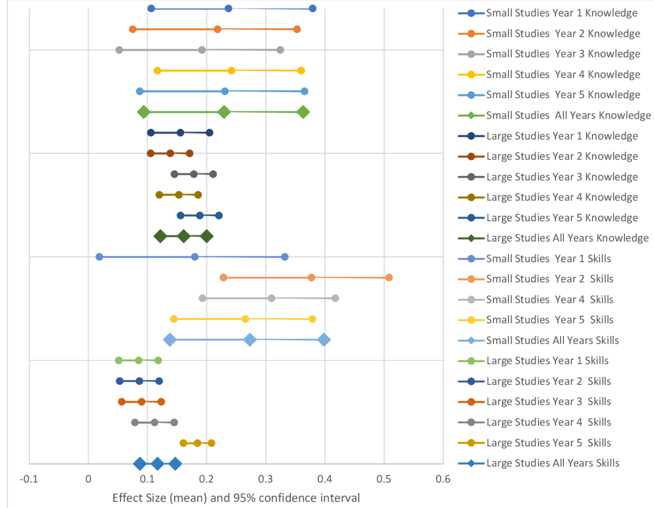

Figure 7 Relationships between UKCAT total score and assessments outcomes by study size. UKCAT, UK Clinical Aptitude Test.

The variety of models used by researchers in regression analyses creates a challenge regarding the presentation of these data. However, these analyses further support findings outlined above that UKCAT has a low but significant relationship with performance outcomes. Effects tended to reduce once other prior achievement was considered.

Incremental validity

Possibly because of the lack of available data and/or the complexity of the analysis, few studies explored the extent to which the UKCAT predicts performance over and above conventional measures of academic attainment such as A-levels. McManus *et al*[9] concluded that the incremental validity of the UKCAT after taking educational attainment into account was 'small but significant'. This finding was further supported by Tiffin *et al*[23].

Predicting other progression outcomes

Tiffin and Paton[24] analysed the odds of students passing at first sitting (compared with other academic outcomes such as fail or resit). The UKCAT SJT score predicted the odds of passing at first attempt (OR=1.28). The authors interpret this finding: 'for every 1 SD above the mean for applicants scored on the UKCAT SJT, the odds of passing first time will increase by around 28%.'

Adam *et al*[14 15] looked at predictors of appearing in the top 20% and bottom 20% of performers at medical school. Being in the top 20% of achievers in year 5 written examinations was associated with having a higher UKCAT total score (higher quantitative reasoning, verbal reasoning and abstract reasoning scores in particular).

The differences in average UKCAT total scores were reported by McManus *et al*[9] for students with different end of year 1 progression markers. The mean average score for students that passed all assessments first time was 2544, for those who passed after resits this was 2486 and those required to repeat the year 2457.

Table 4 Regression analyses

Study	Summary of findings
Adam *et al*[14]	Prior ability (including UKCAT) predicts some year 4 and year 5 outcomes although 'other academic achievement' predicts stronger than UKCAT; demographic variables (gender, age, domicile) reduce the effect of prior achievement.
Husbands and Dowell[26]	UKCAT scores explain 6% of the variance in Dundee year 1 assessments.
Husbands *et al*[7]	UKCAT Total Score explains 6%–13% of variance in Aberdeen Year 4 and 5 exams and 11% of variance in Dundee year 4 examinations.
Hanlon *et al*[22]	No meaningful relationships reported between UKCAT scores and assessments outcomes.
MacKenzie *et al*[13]	UKCAT Total Score and all subtest scores were significantly and positively associated with all four outcome measures in UKFPO applicants.
McManus *et al*[9]	The incremental validity of the UKCAT over current educational attainment was small but significant.
Sartania *et al*[25]	UKCAT total score was independently associated with course performance before and after adjustment for gender, age, ethnicity and deprivation. UKCAT scores predict knowledge outcomes although in most cases, effects reduce on adjustment for the effect of advanced qualifications.
Tiffin and Paton[24]	SJT scores remained significant predictors of theory performance even after adjustment of cognitive ability.
Yates and James[29]	UKCAT total score has a weak relationship with two curriculum themes. Quantitative Reasoning predicts outcomes in one theme and Verbal Reasoning in two themes.
Yates and James[30]	UKCAT total score, Quantitative Reasoning and Verbal Reasoning showed significant effects for a knowledge based exam.
Wright and Bradley[6]	UKCAT a significant predictor of performance in almost all exams.

SJT, situational judgement; UKCAT, UK Clinical Aptitude Test.

The increasing ability of the UKCAT to predict performance as students progressed through their courses (over and above prior attainment) was reported by Tiffin *et al*[23]. Similarly, Sartania *et al*[25] concluded that while both UKCAT and a locally derived science score predicted year 1 performance, UKCAT was the only preadmission measure to independently predict final course performance ranking.

MacKenzie *et al*[8] reported an unexpected relationship between the UKCAT and the UKFPO SJT (r=0.208).

DISCUSSION

Overview

This study for the first time synthesises outcomes from articles reporting on the predictive validity of the UKCAT. Relationships between the UKCAT and medical school outcomes were observed in the vast majority of the 22 identified studies with researchers reporting on these relationships using a range of analyses. Results provide evidence that the UKCAT predicts performance in medical school assessments. Relationships are generally weak although noticeably stronger for both the UKCAT total score and the verbal reasoning subtest. There is some evidence that UKCAT continues to predict performance throughout medical school with some observed upwards trends in prediction over the course of medical programmes.

Main findings

Verbal reasoning scores appear to predict rather more as a stand-alone subtest than the other subtests. It has also been reported that this subtest has the lowest correlation with education attainment (school leaver) of all the subtests[9] suggesting incremental value over (mainly) science school leaver qualifications. This finding ought to be of interest to selectors—if verbal reasoning predicts outcomes rather better than other aspects of the test and measures something different to other performance measures, there may be a case for it being treated differently to the other subtests in selection. Verbal reasoning is also relatively underweighted as part of the total score (having a lower mean average); scaling adjustments could be made to address this if this subtest is deemed to be of greater value.

McManus *et al* reported that the correlation with assessment outcomes for mature applicants (r=0.252, p<0.001) was higher than non-mature (r=0.137, p<0.001). If the UKCAT offers more utility for this group, it allows selectors to compare such applicants who may be offering a diverse range of qualifications (potentially non-standard). There is little subgroup analysis included in these studies—it would be reasonable to speculate that the UKCAT might also predict better for other diverse subgroups of applicants such as international and widening access applicants.

Two of the larger and more detailed studies[9 23] confirm that UKCAT had small incremental validity over and above other measures of academic attainment available at the point of selection. Tiffin *et al*'s findings were across all years of medical programmes which is of particular interest because the paper also reports the declining ability of prior academic achievement to predict outcomes over the course of medical programmes.

Stronger relationships were observed in smaller single and dual centre studies identifying a tension between the power of large cohort studies (with large sample sizes and access to a potentially greater range of consistent demographic markers) against the ability of more local studies to use more fine grained (and arguably more meaningful) outcome markers. The large cohort studies included in this studied were only able to use high level end of year assessments without any direct knowledge of exam content. Understanding the differences between the studies and potential magnitude of differences in outcomes will be helpful when interpreting future studies.

The ability of the UKCAT to predict final course outcomes[23 25] might be explained if the impact of innate cognitive performance continues while the impact of prior education declines. The relationship between the UKCAT SJT and UKFPO results is more difficult to explain, suggesting the need for a greater understanding of the traits being measured in this subtest.

Limitations in identified studies

Outcome markers

Authors comment on the limitations in predictive validity studies created by a lack of relevant outcome markers. Adam *et al*[14] sought to address this issue by looking at a wide range of fine grained outcome markers. Husbands and Dowell[26] also noted the need to further investigate how selection criteria predicted 'specific cognitive and non-cognitive attributes for which they were designed'. The identification of relevant outcome markers has created additional challenges in attempts to validate the UKCAT SJT, with researchers utilising bespoke outcome measures in studies[17] due to the difficulties of identifying existing relevant assessments within medical and dental programmes.

Use of the UKCAT in selection

Universities use the UKCAT in different ways, at different stages of their selection processes and with differing weightings.[1] The actual use of the test in selection may well impact on individual study outcomes. For example, the range of test scores will be more restricted where the test is used as an initial threshold. Universities which use the test in a lighter touch way may well have a greater range of scores represented in the student population.

Interpreting results

The audience for research into admissions is a diverse one. Medical educators are interested in the outcomes alongside others involved in medical selection, applicants and their advisors. In light of this, additional interpretation of findings in studies is desirable. In the studies reviewed, some authors provided additional interpretations to demonstrate the strength of reported relationships.[6 8]

Range restriction

Range restriction creates challenges with these studies because outcomes can only be observed for successful applicants, who are likely as a group to have scored higher in the UKCAT than unsuccessful applicants. Some studies adjusted correlations for range restriction, providing these findings alongside unadjusted outcomes. The adjustments observed by Husbands and Dowell[26] result in stronger correlations between UKCAT and outcome markers (r=0.34 after adjustment vs r=0.25; r=0.24 after adjustment vs r=0.18). Tiffin *et al*[23] included figures which show the magnitude of increases in coefficients following correction for range restriction. The relationships were stronger in all cases once this adjustment had taken place, with the largest increases observed for verbal reasoning and total score.

Wright and Bradley[6] comment on the limitations of restricted range in their work but preferred to report the more conservative approach given the recognised limitations of adjusting.

There has been perhaps a reluctance on the part of researchers to use such methods in case they might be regarded as 'artificially' increasing relationships. It has however been argued that corrected correlations are less biased than those reported without correction.[27]

Implications for the future

UKCAT should encourage stakeholders to continue to undertake predictive validity studies in order to further inform the development of the test and selection processes more generally. The creation of the UKMED will facilitate researchers undertaking full cohort studies (across an entire or multiple medical student intake(s)) with a consistent range of demographic, academic achievement and progression markers. At the same time, however, more local studies able to investigate the ability to predict individual assessment outcomes will continue to have utility. UKCAT should consider supporting a large cohort study every 5 years to ensure that validity evidence remains up to date.

Recommendations for future studies

Those undertaking future studies should:

- Provide consistent detail regarding assessment outcomes being investigated.
- Interpret findings clearly for the benefit of selectors (and test takers).
- Consider performance in the middle years of medical school as fewer studies have looked at years 3 and 4.
- Focus on the lack of evidence regarding dentistry.
- Undertake analysis of relevant subgroup differences (eg, age, gender, international, widening access) with regard to prediction of outcomes.

- Include analysis which adjusts for range restriction or at least comment explicitly on the limiting impact of not undertaking such analysis.

UKMED opens up the opportunity to explore the extent to which all selection criteria (including other aptitude tests) predict performance in medical school and the interaction between such criteria. Future studies will increasingly be able to investigate outcomes beyond medical school into postgraduate training and beyond. This will open up opportunities to investigate how factors used in selection predict career progression and choices.

The work undertaken by Tiffin *et al*[23] in particular would lead naturally to further studies investigating the extent to which the UKCAT might compensate for lower A-level achievement and the impact this might have on opening up routes to widen access.

The UKCAT Consortium should consider the findings from this systematic review in relation to the future development and use of the UKCAT itself. Verbal reasoning clearly plays a great part in the relationships observed. If (as McManus suggests) this subtest also correlates least with prior attainment (A-levels and equivalent) then there is a case to be made for this subtest having a higher weighting or at least being treated differently in selection. A radical approach along these lines would logically increase the ability of the test to predict assessment outcomes in the future.

Strengths and limitations

A particular challenge of longitudinal studies in selection is that during the time required to observe relevant output measures nothing else stands still. This systematic review allows us to draw conclusions from studies over a significant time period. Even so, the UKCAT test itself has gone through significant change as have university curricula; the expansion in student numbers may also impact on applicant demographics and test performance. While we can assume that outcomes reported here might be generalisable to an extent, there is an ongoing need to undertake further studies to reassure future selectors and test takers that the UKCAT remains fit for purpose.

Variability between studies makes generalising across them challenging. Studies took place over a number of years and so year 1 in one school was not the same year 1 (by calendar year) in another. Outcome markers varied in nature and number between schools.

In order to provide greater interpretation of complex data, some results are presented as aggregated outcomes by year group and ought as such be treated with caution.

CONCLUSION

This systematic review supports the use of the UKCAT in selection as the test predicts performance in medical school. However, the relationship is small and selectors ought not to use the test in isolation but alongside other selection criteria.

The UKCAT Consortium should reflect in particular findings regarding verbal reasoning which perhaps support consideration of this subtest separate to the total test score used most frequently by universities.

Further studies are required so that relationships between a changing/developing UKCAT and changing/developing medical school curricula continue to be understood.

Contributors RG is the guarantor. RG undertook the systematic review, identified articles for inclusion, analysed the data, drafted and finalised the manuscript. SN advised on methodology, reviewed and approved outcomes from the systematic review, advised on presentation of results and contributed to drafting of the manuscript. SA undertook an independent secondary review of the search for articles, advised on methodology, advised on presentation of results and contributed to drafting of the manuscript.

Funding The UCAT Consortium supports the cost of RG's PhD and funded the cost of publication of this article (no award number).

Competing interests RG is the (paid) Chief Operating Officer of the UCAT Consortium. SN was an unpaid member of the Board of the UCAT Consortium from 2008 to 2019.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Table with corresponding statistics from included studies is available on request from the corresponding author RG, rachel.greatrix@nottingham.ac.uk.

ORCID iD

Rachel Greatrix http://orcid.org/0000-0002-0341-4079

REFERENCES

1. Greatrix R, Dowell J. UKCAT and medical student selection in the UK - what has changed since 2006? *BMC Med Educ* 2020;20:292.
2. 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.
3. McManus IC, 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;11:242.
4. McManus C, Woolf K, Dacre JE. Even one star at A level could be "too little, too late" for medical student selection. *BMC Med Educ* 2008;8.
5. Lynch B, Mackenzie R, Dowell J, *et al*. Does the UKCAT predict year 1 performance in medical school? *Med Educ* 2009;43:1203–9.
6. Wright SR, Bradley PM. Has the UK clinical aptitude test improved medical student selection? *Med Educ* 2010;44:1069–76.
7. 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.
8. 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.
9. 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 12 UK medical schools. *BMC Med* 2013;11.
10. 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.

11 Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100–28.
12 Wilson D. Practical meta-analysis effect size calculator, 2001. Available: https://campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php
13 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.
14 Adam J, Bore M, Childs R, *et al*. Predictors of professional behaviour and academic outcomes in a UK medical school: a longitudinal cohort study. *Med Teach* 2015;37:868–80.
15 Adam J, Bore M, McKendree J, *et al*. Can personal qualities of medical students predict in-course examination success and professional behaviour? an exploratory prospective cohort study. *BMC Med Educ* 2012;12:69.
16 McAndrew R, Ellis J, Valentine RA. Does a selection interview predict year 1 performance in dental school? *Eur J Dent Educ* 2017;21:108–12.
17 Patterson F, Cousans F, Edwards H, *et al*. The predictive validity of a Text-Based situational judgment test in undergraduate medical and dental school admissions. *Acad Med* 2017;92:1250–3.
18 Lala R, Wood D, Baker S. Validity of the UKCAT in applicant selection and predicting exam performance in UK dental students. *J Dent Educ* 2013;77:1159–70.
19 Lambe P, Kay E, Bristow D. Exploring uses of the UK clinical aptitude Test-situational judgement test in a dental student selection process. *Eur J Dent Educ* 2018;22:23–9.
20 Lambe Paul BD. Can a situational judgement test replace a standardised interview as part of the process to select undergrdauate medical students?, in OTTAWA2016: Australia 2016.
21 Bowman NA. Effect sizes and statistical methods for meta-analysis in higher education. *Res High Educ* 2012;53:375–82.
22 Hanlon K, Prescott G, Cleland J, *et al*. Does UKCAT predict performance in the first year of an integrated systems based medical school curriculum? *Medical Education* 2011;2:87.
23 Tiffin PA, Mwandigha LM, Paton LW, *et al*. Predictive validity of the UKCAT for medical school undergraduate performance: a national prospective cohort study. *BMC Med* 2016;14:140.
24 Tiffin PA, Paton LW. *Exploring the validity of the 2013 UKCAT SJT-prediction of undergraduate performance in the first year of medical school: summary version report*. Nottingham: UKCAT, 2017.
25 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.
26 Husbands A, Dowell J. Predictive validity of the Dundee multiple mini-interview. *Med Educ* 2013;47:717–25.
27 Zimmermann S, Klusmann D, Hampe W. Correcting the predictive validity of a selection test for the effect of indirect range restriction. *BMC Med Educ* 2017;17:246.
28 Foley JI, Hijazi K. Predictive value of the admissions process and the UK clinical aptitude test in a graduate-entry dental school. *Br Dent J* 2015;218:687–9.
29 Yates J, James D. *The value of the UK Clinical Aptitude Test in predicting pre-clinical performance: a prospective cohort study at* Nottingham Medical School. *BMC Med Educ* 2010;10:55.
30 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.
31 Mwandigha LM, Tiffin PA, Paton LW, *et al*. What is the effect of secondary (high) schooling on subsequent medical school performance? a national, UK-based, cohort study. *BMJ Open* 2018;8:e020291.
32 Srikathirkamanathan KCS, MacManus BN. The relationship between UKCAT scores and finals exam performance for widening access and traditional entry students 2018.