

Entry tests for graduate medical programs: is it time to re-think?

Michele A Groves, Jill Gordon and Greg Ryan

Admission to medical school is highly competitive and generally based on a combination of academic achievement, written tests and an interview. In Australia, two tests are widely used: the Graduate Australian Medical School Admissions Test (GAMSAT) and the Undergraduate Medicine and Health Sciences Admission Test (UMAT). GAMSAT is designed to assess problem-solving and data interpretation in the social, physical and biological sciences, as well as critical thinking, reasoning and written communication.¹ In 2005, GAMSAT contributed to the selection of about half the medical students in Australia. However, no systematic analysis of GAMSAT's predictive validity has been published.

The predictive validity of admissions tests is contentious, with inconsistent results from studies evaluating the reliability of cognitive measures as indicators of academic or clinical performance.²⁻⁴ In the United States, Mitchell et al found that the Medical College Admission Test (MCAT)-predicted grades in medical school were only slightly better than performance at school.⁵ In contrast, a review of medical schools in the United Kingdom concluded that pure measures of reasoning ability are less predictive than measures of knowledge, such as A-levels.⁶ However, there is evidence that personality factors have been underexplored and that situational tests which measure particular constructs may be more useful than cognitive tests.⁷ Interviews, another popular measure, may in fact be poor predictors of behavioural and attitudinal aptitude. Although interviews attempt to assess personality characteristics, their inherently subjective nature may reduce reliability.⁸

A hallmark of expert clinical performance is skill in clinical reasoning. Although it has been shown that clinical reasoning develops throughout undergraduate medical training,⁹ studies investigating the predictive validity of students' entry characteristics commonly focus on their relationship with overall academic performance or performance as interns.¹⁰⁻¹³ Two studies have investigated the association between pre-admission test scores and clinical performance, including clinical reasoning, at graduation,^{14,15} and they yielded conflicting results.

We sought to confirm and extend the findings of a previous study at one graduate-entry medical school.¹⁶ That study found a

ABSTRACT

Objective: To examine the relationship between medical school applicants' performances in the Graduate Australian Medical School Admissions Test (GAMSAT) and structured interviews and their subsequent performance in medical school.

Design: Students in Years 2–4 of two graduate-entry medical programs were invited to complete two previously validated tests of clinical reasoning. These results and their Year 2 examination results were compared with their previous performance in GAMSAT and at interview.

Setting: The graduate-entry programs at the Universities of Queensland and Sydney.

Participants: 189 student volunteers (13.6% response rate).

Main outcome measures: Students' test results on a set of Clinical Reasoning Problems (CRPs) and a Diagnostic Thinking Inventory (DTI) and their Year 2 examination results.

Results: There was no association between performance in GAMSAT and performance in the CRPs; there was a weak negative correlation between performance in GAMSAT and the DTI ($-0.05 > r > -0.31$, $P = 0.03$). The correlation between GAMSAT and examination results was weak ($r < 0.24$, $P = 0.02$). The correlation between GAMSAT and interview scores for each school was weakly negative for University of Queensland ($r = -0.34$, $P < 0.01$) and weakly positive for University of Sydney ($r = 0.11$), with a combined significance level $P < 0.01$.

Conclusions: We did not find evidence that GAMSAT and structured interviews are good predictors of performance in medical school. Our study highlights a need for more rigorous evaluation of Australian medical school admissions tests.

MJA 2007; 186: 120–123

For editorial comment, see page 118

significant relationship between pre-admission interview performance and development of clinical reasoning skill, but no relationship between interview performance and GAMSAT results or between GAMSAT results and the development of clinical reasoning skill in medical school. In this study, we asked to what extent do students' entry characteristics, particularly performance against specific admission criteria, predict clinical reasoning skill and academic performance?

METHODS

Participants

Participants were students enrolled in Years 2–4 of the graduate-entry medical programs at the Universities of Queensland (UQ) and Sydney (USyd), recruited on a voluntary basis.

Admissions criteria: GAMSAT and interview

GAMSAT consists of three sections: Reasoning in Humanities and Social Sciences, Written Communication, and Reasoning in Biological and Physical Sciences. It has been designed to:

assess the capacity to undertake high level intellectual studies in a demanding course... GAMSAT evaluates the nature and extent of abilities and skills gained through prior experience and learning, including the mastery and use of concepts in basic science as well as the acquisition of more general skills in problem solving, critical thinking and writing.¹⁷

Both UQ and USyd offer a semi-structured interview to all applicants who have met a threshold GAMSAT score. The interview is designed to evaluate communication skills, cognitive style and decision-making ability, cooperativeness and participation, motivation, and personal attributes, including empathy and self-awareness.

At UQ, enrolment offers are based equally on GAMSAT and interview score, while at USyd, initial offers are based on interview score only.

Measurement of clinical reasoning

Two instruments were used to assess clinical reasoning: a set of Clinical Reasoning Prob-

lems (CRPs),¹⁸ and the Diagnostic Thinking Inventory (DTI).¹⁹ Although both are reliable and valid methods of assessing clinical reasoning,¹⁸ their emphases are different. The CRPs focus on the integration of knowledge (biomedical and clinical) with reasoning ability, while the DTI is a self-report questionnaire that explores the respondent's cognitive approach to the clinical reasoning process, and is independent of knowledge. The DTI has been widely used to investigate the relationship between clinical reasoning skill and various student characteristics such as learning style, communication skill and preclinical background.²⁰⁻²³ Both instruments were used in the original study.¹⁶

Test procedure

Participants completed the DTI and the CRPs without supervision and without a time limit. Participants were asked not to refer to textbooks or consult colleagues.

The entry characteristics (sex, age at commencement of the program, nature of previous tertiary qualification, and scores on selection criteria [GAMSAT and interview]) for participants and non-participants were provided on request from each school and recorded, along with year of medical training and academic performance. For participants, second-year examination results were obtained and used as the measure of academic performance, as this is a common assessment point for both universities.

Statistical analysis

We used the same analytical approach as in the original work.¹⁶ For each medical school, descriptive statistics, contingency tables and *t* tests were performed for participants and their corresponding year group to determine whether the results derived from the study data were generalisable to the entire year group.

We evaluated the predictive power of entry characteristics for clinical reasoning skill and academic performance by the following method. Categorical variables were age (17–24 years and ≥ 25 years), sex and academic background (biological or non-biological science degree). Outcome variables were total CRP score, DTI scores (flexibility, structure and total), and Year 2 exam scores. We correlated these variables with GAMSAT score and interview score. One-way analysis of variance was used to assess the relationship between these variables and student entry characteristics (cohort, age,

sex, and academic background). We then used general linear modelling to determine any relationship between academic performance and characteristics that had a significant association with clinical reasoning scores. The final model was selected by backwards elimination until only significant terms remained, and then screened for interactions between the terms.

To maximise statistical power, Fisher's method was used to calculate combined significance levels for each outcome variable.

RESULTS

Participation

A total of 189 students participated over the 2003 academic year (Box 1), representing an overall participation rate of 13.6%.

Comparison of participants and non-participants within each year group

For both universities, no significant differences were found between participants and non-participants in their year group with regard to sex, academic background, GAMSAT or interview score. However, Year 3 participants were older than non-participants at both UQ ($P < 0.01$) and USyd ($P = 0.04$). Participating Year 3 students at UQ also scored higher on the written communication component of GAMSAT than their year group (69.5 compared with 64.1, $P < 0.01$). At USyd, Year 2 students who participated scored higher on this component of GAMSAT than their year group (65.9 compared with 63.4, $P = 0.02$).

Comparison of UQ and USyd participants

UQ participants were generally younger and scored higher on GAMSAT, including two of the three subscales, than USyd participants. There were no differences in interview scores. Both groups showed a progressive increase in scores on the two clinical reasoning indicators (CRPs and DTI). In each case, the structure of knowledge subscale of the DTI was primarily responsible for this increase.

1 Student numbers and participation rates

Year	University of Queensland			University of Sydney		
	N _{Participants}	N _{Cohort}	Rate (%)	N _{Participants}	N _{Cohort}	Rate (%)
2	29	238	12.2%	50	227	22.0%
3	18	239	7.5%	30	234	12.8%
4	34	229	14.8%	28	221	12.7%
Total	81	706	11.5%	108	682	15.8%

Predictors of clinical reasoning and academic performance

Analysis of three of the five entry characteristics (age, sex and academic background) found no age- or sex-related differences for clinical reasoning or academic performance. However, students with a non-biological science-based primary degree scored lower for both the CRPs ($P < 0.01$) and the Year 2

exams ($P = 0.03$). No differences were found in relation to DTI scores (Box 2).

Box 3 shows the correlation of the remaining characteristics, GAMSAT and interview scores, with scores for the CRPs, DTI and the second-year barrier examination. There was a weak negative correlation between scores for GAMSAT (total) and the DTI, which was significant when both universities' data were combined ($-0.05 > r > -0.31$, $P = 0.03$), but no association with the CRPs. The correlation between GAMSAT and examination results was weak ($r < 0.24$, $P = 0.02$). The correlation between GAMSAT and interview scores was weakly negative for UQ ($r = -0.34$, $P < 0.01$) and weakly positive for USyd ($r = 0.11$), with a combined $P < 0.01$. Neither school showed an association between interview scores and any of the outcome variables.

General linear modelling showed no interdependence between entry characteristics in relation to their ability to predict scores on the CRPs, DTI or Year 2 examinations.

DISCUSSION

We found little evidence that students' entry characteristics significantly predict performance during medical training. The minor differences in average entry characteristics between the two universities generally reflect differences in admission criteria or the secondary education system; for example, schools in New South Wales offer an extra year of pre-university education compared with Queensland, and applicants to the USyd's medical program require a minimum grade point average (GPA) of 5, compared with 4.5 for UQ. These differences have a small effect on the threshold GAMSAT scores at each school.

We found no age- or sex-related differences in clinical reasoning or academic performance. Although a non-biological science-based primary degree was shown to have an impact, this is not large enough to cause a difference in pass rate, and in-house analysis of assessment data at UQ has consistently indicated that any differences between students based on academic back-

2 Relationship of clinical reasoning and academic achievement to medical school entry characteristics*

	CRPs	DTI flexibility	DTI structure	DTI total	Year 2 score
University of Queensland (n = 81)					
Sex					
Male	179.10 (30.97)	88.28 (7.92)	81.54 (9.72)	169.82 (15.12)	69.77 (5.82)
Female	190.21 (31.86)	89.88 (8.46)	83.05 (9.26)	173.24 (15.37)	69.30 (6.36)
P	0.12	0.38	0.48	0.32	0.74
Age					
17–24 years	187.26 (31.98)	89.48 (8.37)	83.02 (9.52)	172.71 (15.09)	69.79 (5.87)
≥ 25 years	177.05 (30.42)	87.89 (7.67)	80.05 (9.11)	167.95 (15.58)	68.59 (6.83)
P	0.22	0.46	0.23	0.24	0.48
Academic background					
Biological	189.30 (31.47)	89.48 (8.60)	82.70 (9.52)	172.38 (15.51)	70.19 (6.04)
Non-biological	167.00 (25.67)	87.07 (6.07)	79.71 (8.85)	166.79 (13.46)	66.47 (5.37)
P	0.02	0.32	0.28	0.21	0.04
University of Sydney (n = 108)					
Sex					
Male	171.02 (32.24)	89.93 (9.09)	83.05 (7.85)	173.02 (14.36)	0.27 (0.76)
Female	178.82 (36.03)	85.68 (9.68)	81.54 (10.19)	167.06 (18.13)	0.40 (0.80)
P	0.26	0.03	0.42	0.08	0.38
Age					
17–24 years	177.20 (38.67)	87.69 (9.54)	82.39 (9.77)	170.08 (17.37)	0.36 (0.77)
≥ 25 years	173.91 (28.25)	86.73 (9.88)	81.70 (8.82)	168.23 (16.52)	0.34 (0.81)
P	0.63	0.61	0.71	0.58	0.87
Academic background					
Biological	180.35 (35.22)	86.86 (9.77)	82.61 (10.03)	169.48 (17.68)	0.42 (0.82)
Non-biological	163.62 (30.54)	88.48 (9.36)	80.76 (7.21)	168.90 (15.16)	0.17 (0.66)
P	0.03	0.44	0.37	0.87	0.14
University of Queensland and University of Sydney combined P values (n = 189)					
Sex	0.14	0.06	0.52	0.14	0.64
Age	0.41	0.64	0.46	0.41	0.78
Academic background	< 0.01	0.42	0.34	0.49	0.03

* Values for scores are mean (SD). CRPs = Clinical Reasoning Problems. DTI = Diagnostic Thinking Inventory. ♦

ground have disappeared by graduation. At USyd, there is a whole-year assessment in Year 3, but no final year assessment. However, it seems likely from the available evidence that we would not find differences at the time of graduation. Nor did we find any evidence to suggest that academic background contributes to the correlation of GAMSAT scores with Year 2 results.

Our results show that students from non-biological science-based academic backgrounds are not seriously disadvantaged either by the selection process or during the program; medical schools that actively seek a more diverse intake can be reassured that these students are just as likely to succeed.

The poor correlation between GAMSAT and interview scores suggests that they do assess different constructs, as intended. However, we found little evidence that either predicts clinical reasoning or academic performance to any substantial extent. Indeed, GAMSAT's negative correlation with the DTI, a knowledge-independent measure of reasoning, and its small positive correlation with examination scores reinforce the argument that GAMSAT is primarily a measure of knowledge. Furthermore, we were unable to confirm previous findings of a significant relationship between pre-admission interview and subsequent clinical reasoning skill.¹⁶ Although this may be due to differences in the interview questions between

the two universities, it could simply reflect differences in the sample size and statistical power of the two studies. Nevertheless, our findings suggest that interviews do not discriminate well between applicants,⁸ and support recent moves to other formats that are less subjective and produce more reliable and valid outcomes.^{24–26}

A limitation of our study is that medical students represent a relatively small sub-group of all applicants to medicine, namely those who achieved the required GAMSAT threshold score. Nevertheless, the fact that GAMSAT appears to be a negative predictor of clinical reasoning is a cause for concern, given that it was specifically designed to assess reasoning in both the social and biological sciences.

The combined student response rate of 13.6% allows for sampling error and affects the study's confidence level. Although detailed statistical analysis was beyond the scope of this study, our findings do confirm and extend the results of previous research¹⁶ with regard to GAMSAT.

Most medical schools try to select students who will develop into safe and effective medical practitioners. Research in this area usually proceeds in two steps: (1) do entrance exams predict performance in medical school? and (2) does performance in medical school predict performance in practice? We did not find a strong association between performance in the two most widely used selection measures and in medical school. Other studies have failed to find a strong association between performance in medical school and clinical performance after graduation. Medical school selection is such a high-stakes process that we need to clearly define the performance indicators, including such measures as intern assessments, to be targeted by the selection process. We then need to evaluate both GAMSAT and the interview against these indicators. Without such evidence, it is difficult to resist arguments for a process that is less expensive, elaborate, stressful and time-consuming, such as the GPA-weighted lottery method of selection used in the Netherlands.²⁷

CONCLUSION

Despite the limitations of our study, the results challenge the predictive validity of GAMSAT and the reliability of admission interviews, even when they are semi-structured. Because entry to medical school is a highly valued goal and a well educated medical workforce such an important national asset, there is an urgent need for collaborative studies to explore these issues further. Many new medical schools are in development or recently accredited and

3 Correlation of medical school admissions criteria (GAMSAT, interview) with clinical reasoning (CRPs, DTI) and academic achievement (Year 2 results)

	CRPs		DTI Flexibility		DTI Structure		DTI Total		Year 2 results	
	r	P	r	P	r	P	r	P	r	P
University of Queensland (n = 80)										
Total GAMSAT	0.10	0.93	-0.28	0.01	-0.28	0.01	-0.31	<0.01	0.23	0.05
Humanities and social sciences	0.11	0.34	-0.18	0.12	-0.12	0.31	-0.15	0.18	0.06	0.60
Written communication	-0.04	0.75	0.01	0.94	-0.07	0.53	-0.04	0.70	-0.03	0.81
Biological and physical sciences	0.01	0.91	-0.23	0.04	-0.22	0.05	-0.24	0.03	0.21	0.06
Interview	-0.02	0.88	0.11	0.33	0.16	0.15	0.16	0.17	-0.19	0.09
University of Sydney (n = 108)										
Total GAMSAT	0.04	0.69	-0.04	0.72	-0.05	0.59	-0.05	0.66	0.18	0.07
Humanities and social sciences	-0.04	0.70	-0.14	0.17	-0.09	0.38	-0.13	0.20	0.13	0.20
Written communication	-0.18	0.08	0.06	0.55	-0.02	0.82	0.03	0.78	-0.02	0.82
Biological and physical sciences	0.14	0.16	-0.01	0.89	-0.03	0.74	-0.02	0.83	0.20	0.04
Interview*	-0.05	0.60	-0.10	0.30	-0.14	0.16	-0.13	0.20	-0.01	0.91
University of Queensland and University of Sydney combined (n = 188)										
Total GAMSAT	—	0.93	—	0.04	—	0.04	—	0.03	—	0.02
Humanities and social sciences	—	0.58	—	0.10	—	0.37	—	0.16	—	0.37
Written communication	—	0.23	—	0.86	—	0.80	—	0.88	—	0.94
Biological and physical sciences	—	0.43	—	0.15	—	0.16	—	0.12	—	0.02
Interview*	—	0.87	—	0.33	—	0.11	—	0.15	—	0.29

*Represents adjusted score for USyd participants. CRPs = Clinical Reasoning Problems. DTI = Diagnostic Thinking Inventory. GAMSAT = Graduate Australian Medical School Admissions Test.

would benefit from the experience of more established schools in avoiding error.

COMPETING INTERESTS

None identified.

AUTHOR DETAILS

Michele A Groves, BSc, PhD, Deputy Head, Director of Medical Studies¹

Jill Gordon, MB BS, PhD, Director, Centre for Medical Humanities²

Greg Ryan, RN, PhD, Director, Pharmacy Education Unit, Faculty of Pharmacy²

¹ School of Medicine, Griffith University, Gold Coast, QLD.

² University of Sydney, Sydney, NSW.

Correspondence: m.groves@griffith.edu.au

REFERENCES

- Aldous CJH, Leeder SR, Price J, et al. A selection test for Australian graduate-entry medical schools. *Med J Aust* 1997; 166: 247-250.

- Salvatori P. Reliability and validity of admissions tools used to select students for the health professions. *Adv Health Sci Educ Theory Pract* 2001; 6: 159-175.

- Kulatunga-Moruzi C, Norman G. Validity of admissions measures in predicting performance outcomes: the contribution of cognitive and non-cognitive dimensions. *Teach Learn Med* 2002; 14: 34-42.

- Julian ER. Validity of the Medical College Admission Test for predicting medical school performance. *Acad Med* 2005; 80: 910-917.

- Mitchell K, Haynes R, Koenig J. Assessing the validity of the updated Medical College Admission Test. *Acad Med* 1994; 69: 394-401.

- 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: 555-559.

- Lievens F, Coetsier P. Situational tests in student selection: an examination of predictive validity, adverse impact and construct validity. *Int J Selection Assess* 2002; 10: 245-257.

- Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ* 2004; 38: 314-326.

- Groves M. The clinical reasoning process: a study of its development in medical students [PhD thesis]. Brisbane: University of Queensland, 2002.

- Blue AV, Gilbert GE, Elam CL, Basco WT Jr. Does institutional selectivity aid in the prediction of medical school performance? *Acad Med* 2000; 75 (10 Suppl): S31-S33.

- de Clercq L, Pearson S, Rolfe I. The relationship between previous tertiary education and course performance in first year medical students at Newcastle University, Australia. *Educ Health (Abingdon)* 2001; 14: 417-426.

- Grey M, Pearson S, Rolfe I, et al. How do Australian doctors with different pre-medical school backgrounds perform as interns? *Educ Health (Abingdon)* 2001; 14: 87-96.

- Veloski JJ, Callahan CA, Xu G, et al. Prediction of students' performances on licensing examinations using age, race, sex, undergraduate GPAs, and MCAT scores. *Acad Med* 2000; 75 (10 Suppl): S28-S30.

- Basco WT Jr, Gilbert GE, Chessman AW, Blue AV. The ability of a medical school admission process to predict clinical performance and patients' satisfaction. *Acad Med* 2000; 75: 743-747.

- Vu NV, Dawson-Saunders B, Barrows HS. Use of a medical reasoning aptitude test to help predict performance in medical school. *J Med Educ* 1987; 62: 325-335.

- Groves M, O'Rourke P, Alexander H. The association between student characteristics and the development of clinical reasoning in a graduate-entry, PBL medical programme. *Med Teach* 2003; 25: 626-631.

- Australian Council for Educational Research. GAMSAT Graduate Australian Medical School Admissions Test information booklet 2007. Melbourne: ACER, 2006. <http://www.gamsat.acer.edu.au/images/infobook/GAMSATInfoBook07.pdf> (accessed Nov 2006).

- Groves M, Scott I, Alexander H. Assessing clinical reasoning: a method to monitor its development in a PBL curriculum. *Med Teach* 2002; 24: 507-515.

- Bordage G, Grant J, Marsden P. Quantitative assessment of diagnostic ability. *Med Educ* 1990; 24: 413-425.

- Round AP. Teaching clinical reasoning — a preliminary controlled study. *Med Educ* 1999; 33: 480-483.

- Sobral D. Appraisal of medical students' diagnostic ability in relation to their learning achievement and self-confidence as a learner. *Med Teach* 2000; 22: 59-63.

- Sobral DT. Diagnostic ability of medical students in relation to their learning characteristics and preclinical background. *Med Educ* 1995; 29: 278-282.

- Windish DM, Price EG, Clever SL, et al. Teaching medical students the important connection between communication and clinical reasoning. *J Gen Intern Med* 2005; 20: 1108-1113.

- Hanafi Z, Abd Hamid M, Kifli N, et al. The use of the multiple mini interview to select medical students. Association of Medical Education in Europe Annual Conference; 2005 Aug 30 – Sep 2; Amsterdam, The Netherlands.

- Lumsden MA, Bore M, Millar K, et al. Assessment of personal qualities in relation to admission to medical school. *Med Educ* 2005; 39: 258-265.

- Powis DA. Selecting medical students: the significance of personal attributes other than academic ability. Newcastle, NSW: Newcastle University, 2005.

- ten Cate TJ, Hendrix HL. [Initial experience with selection procedures for admission to medical school] [Dutch]. *Ned Tijdschr Geneesk* 2001; 145: 1364-1368.

(Received 8 Jun 2006, accepted 9 Nov 2006)