BMJ Open

IS THE ADMISSION TEST FOR THE DEGREE COURSE IN MEDICINE A GOOD PREDICTOR OF MEDICAL STUDENT ACAMEDIC PERFORMANCE? AN ITALIAN EXPERIENCE

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017417
Article Type:	Research
Date Submitted by the Author:	10-May-2017
Complete List of Authors:	Migliaretti, Giuseppe; University of Turin, Clinical and Biological Sciences Bozzaro, Salvatore; Universita degli Studi di Torino Scuola di Medicina, University of Turin, Clinical and Biological Sciences Regione Gonzole, San Luigi Gonzaga Hospital, Orbassano (TO) Turin Siliquini, Roberta; Universita degli Studi di Torino Costa, Giuseppe; Universita degli Studi di Torino Scuola di Medicina Cavallo, Franco; Universita degli Studi di Torino Scuola di Medicina
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Epidemiology, Research methods
Keywords:	Admission test, Degree course in Medicine, academic performances, Predictive values, ROC curve, diagnostic test validity

IS THE ADMISSION TEST FOR THE DEGREE COURSE IN MEDICINE A GOOD PREDICTOR OF MEDICAL STUDENT ACAMEDIC PERFORMANCE? AN ITALIAN EXPERIENCE

Migliaretti Giuseppe, Bozzaro Salvatore, Siliquini Roberta, Costa Giuseppe, Cavallo Franco

School of Medicine,University of Turin

Corresponding author

Giuseppe Migliaretti, School of Medicine University of Turin

Str. Santena 5bis, 10100 Turin (Italy)

Mail: giuseppe.migliaretti@unito.it

Phone: +39 11 6705808

Fax: +39 11 2365808

Co-Authors

Salvatore Bozzaro - School of Medicine University of Turin (ITALY)

Roberta Siliquini - School of Medicine University of Turin (ITALY)

Giuseppe Costa - School of Medicine University of Turin (ITALY)

Franco Cavallo - School of Medicine University of Turin (ITALY)

-Key word: Admission test, Degree course in Medicine, academic performances, diagnostic test, Predictive values, ROC curve, University of turin

-Additional data: No additional data avaible

-Ethical aspects: Work has respected all ethical aspects and has not benefited from any kind of economic support

WORD COUNT: 1784

ABSTRACT

Objectives

The usefulness of university admission tests in School of Medicine has been the subject of much discussion in recent years. In the academic year 2014/2015, many students who were not admitted appealed and won their cases in the regional administrative court, thereby obtaining the right to enrol on degree courses despite not having passed the admission test. This situation has made it possible to evaluate the predictive capacity of admission tests on academic performances in a methodologically consistent sound manner. The aim of the present work is to discuss the preliminary results directed at evaluating the capacity of the admission test for the degree course in medicine to predict the subsequent academic success.

Setting and Partecipant

The present study considers the 683 students who enrolled onto the first year of the degree course in medicine in the academic year 2014/2015 at the University of Turin at the Molinette and San Luigi Gonzaga colleges. The students were separated into two categories: those who passed the admission test (*Regular*) [n1=531] and those who did not pass the admission test but won their appeal in the TAR (*TAR*) [n2=152].

Outcomes

Validity of the admission test was analysed using Positive Predictive values and the ROC curve

Results

The results show that the admission test appears to be a good tool for predicting the academic performances in the first year of course (AUC= 0.69 AUC_95% CI from 0.63 to 0.73); we can also see that some subject areas seem to have a greater discriminating capacity than others: (i.e. Biology).

Conclusions

The initial results of our study seem to confirm the capacity of the admission test to predict the academic success of the students studying on the degree course in medicine, at least with regard to the first year of course of School of Medicine.

Strengths and limitations of this study

-It is among the first research, based on statistically sound methods, conducted in the Italian reality to evaluate the reliability of the admission test

- Further investigations will be required in order to analyse the two cohorts over a longer period of time (at least until the end of the second or third year, when the subject areas no longer reflect those that make up the admission test) and to compare the results obtained with those related to other Italian medical schools

Bullet points

- The study show the capacity of the admission test for the degree course in medicine to predict the subsequent academic performances

-In the study students who passed the admission test (*Regular)* and those who did not pass the admission test but won their appeal in the '*Tribunale Amministrativo Regionale*: TAR' (*TAR*) were compared

- Statistical analysis based on ROC curve, Sensibility of test (Se), Positive Predictive Values (PPV), Negative Predictive Values (NPV), Positive Likelihood ratio (LH+), Negative Likelihood ratio (LH-) and relative 95% Confidence Intervals were performed

- The results show that the admission test appears to be a good tool for predicting those that will be able to reach academic success

- In conclusion, the results seem to highlight the importance of employing an evaluation instrument that permits students to evaluate their potential to undertake a certain degree course

Introduction

The usefulness of university admission tests has been the subject of much discussion in recent years, and particular attention has been placed on the admission tests for degree courses in medicine. In Italy, a proposal was made to eliminate the test and instead make it possible to block the progression of students who did not achieve the minimal grades at the end of their first year, but as yet no decision has been taken to modify the system. Nonetheless, an important issue that still requires further investigation is the extent to which the results of the admission tests are able to predict the future ability of a student to cope with the curriculum. Many studies are present in the literature that highlight a tight relationship between admission test results and academic performance [1, 2] and some reflections have been proposed relative to possible re-evaluations of the assessment process [3]; so far, the large majority of these studies have looked at situations outside Italy [4-19]. One reason for this is the lack of useful data for evaluating the predictive capacity of the admission test. In fact, until the start of the academic year 2013/2014, the admission test scores and data regarding the examinations results, could only be analysed for the students who had passed the admission test. Some studies, comparing Italian Universities, have investigated whether admission test scores are able to predict good first year exam results. Although interesting, the results have not made any a significant contribution to the literature due to the lack of adequate control groups. In the academic year 2014/2015, many students who were not admitted to their desired courses (due to failure to pass the admission test) appealed and won their cases in the regional administrative court (*Tribunale Amministrativo Regionale*: TAR), thereby obtaining the right to enrol on their respective degree courses despite not having passed the admission test. This situation has made it possible to evaluate the predictive capacity of admission tests on students' subsequent academic exam results in a more precise and methodologically sound manner. Within such a framework, the two constituent colleges of the University of Turin (Molinette and San Luigi Gonzaga) have been dedicated to developing this line of research.

Objectives

The aim of the present work is to discuss the preliminary results directed at evaluating the capacity of the admission test for the degree course in medicine to predict the subsequent academic success of students enrolled on this degree course at the two constituent colleges of the University of Turin.

Data base

The present study considers the 683 students who enrolled onto the first year of the degree course in medicine in the academic year 2014/2015 at the University of Turin at the Molinette and San Luigi Gonzaga colleges. The students were divided into two categories: those who passed the admission test (*Regular*) [n1=531] and those who did not pass the admission test but won their appeal in the TAR (*TAR*) [n2=152].

All students were monitored over the period spanning January 2015 to May 2016 (the period over which the first-year exams are taken in the two colleges considered).

Methods

The following data are presented as means, standard deviations (SD), median and 95% confidence intervals for the two investigated student groups ('Regular' and 'TAR'): admission test score; secondary school final grade; number of CFU (*university exam credits*) acquired; and the average first-year exam grade. Mean data were then compared using the non parametric Mann-Whitney test.

Considering as the principal end-point measure the number of CFU accumulated in the observation period, the students were classified into the following categories defining two different Gold Standard reference:

Gold Standard - GS

- Students who acquired less than half of the required credits at the end of the first year (*N*)
- Students who acquired more than half of the required credits at the end of the first year (*P*)

Gold standard - GS2

- Students who failed to obtain some first year CFU (N2)
- Students who acquired all the CFU required for the first year of the course (*P2*)

For both Gold standard reference, independent of the student categories (Regular / TAR), the predictive capacity of the admission test was evaluated using a ROC curve estimation and the relative area under the curve (AUC), with its 95% confidence interval (AUC_IC95%).

In order to assess the weight that each of the subject areas of the admission test had on the predictive capacity of the test, the same analysis was also conducted on the specific score achieved in each discipline.

Finally GS and GS2 were compared with the student categories (Regular / TAR) using Positive Predictive Values (PPV), Negative Predictive Values (NPV), Positive Likelihood ratio (LH+) and Negative Likelihood ratio (LH-); estimates are shown with the relative 95% confidence interval.

Results

Table 1 shows the number of students in each of the two groups (Regular and TAR) enrolling onto the first year of the degree course in medicine at the two University of Turin colleges.

Comparing the two groups with respect to admission test score, secondary school final grade, CFU acquired and average grade achieved in the first-year exams, statistically significant differences were consistently seen (Table 1.1 and Table 2). In particular, with regard to the number of CFU acquired, the results of the admission test appear to be able to identify with a higher probability the students that, on average, obtained better results at the end of the first year.

Independent of student category (Regular or TAR), the predictive capacity of the admission test regarding the GS (*Gold Standard reference*) was assessed using the ROC curve estimate (Figure 1). The results show that the admission test appears to be a good tool for predicting those that will be able to acquire more than half of the required credits CFU, presenting an AUC= 0.69 (AUC_95% CI from 0.63 to 0.73). Regarding the specific subject areas that make up the admission test (Figure 1.1), we can also see that some subject areas seem to have a greater discriminating capacity than others: biology (AUC=0.653, 95% CI from 0.59 to 0.71) and chemistry (AUC=0.654, 95% CI from 0.60 to 0.71) stand out as two such subject areas .

The classification of students as 'Regular' or 'TAR' (indicating the effective outcome of the admission test) was compared with the Gold Standard reference (GS). The results present a high value of Sensitivity [Se = 0.82(95%CI from 0.79 to 0.86)], which contrasts, however, with a low value of specificity (Table 3), thereby highlighting how some students although classified as Regular are still unable to achieve the objectives of the first year's curriculum. These results are clearly conditioned by the different sample numbers of the two groups. The predictive values [PPV and NPV], on the other hand, are much more informative. The results reveal a good predictive ability of the test when the outcome is positive [PPV = 0.79 (95%CI from 0.74 to 0.82)], but relatively low predictive power when it is negative.

The estimate of positive Likelihood LH+ is equal to 1.39 (95% CI from 1.18 to 1.61). This result takes on even greater importance if we consider the estimate of negative likelihood LH-, equal to 0.47 (95% CI from 0.26 to 0.68), which implies that if a student belongs to the TAR group, the probability that they will not achieve half or more of the CFU required at the end of the first year is even greater.

By increasing the reference cut off level, however, we can see that the result of the admission test becomes more predictive by including those that are unable to achieve the first year objectives. In order to assess this aspect, we defined a new Gold Standard reference – Gold Standard 2 (GS2): Students who acquired all the CFU required for the first year of the course (*P2*) vs. Students who failed to obtain some first year CFU (N) (Table 3). Thus, by comparing GS2 with the effective outcome of the admission test ('Regular' vs. 'TAR'), we observe an improvement in the negative predictive capacity of the test [NPV=0.80 (95%CI from 0.73 to 0.86], but a reduction in its positive predictive capacity PPV. However, it must be highlighted that the values of LH do not vary to any great extent, once again confirming the good predictive capacity of the test.
These results indicate that even some 'Regular' students may not be able to achieve the course objective of the first year, and that there is a good probability that those classified as 'TAR' will not achieve it.

Conclusions

The present study was a preliminary explorative experiment limited to the context of the University of Turin. Further investigations will be required in order to analyse the two cohorts over a longer period of time (at least until the end of the second or third year, when the subject areas no longer reflect those that make up the admission test) and to compare the results obtained with those related to other Italian medical schools. Nevertheless, these initial results seem to confirm the capacity of the admission test to predict the academic success of the students studying on the degree course in medicine, at least with regard to the first year of study. It also provides further confirmation of the fact that students with very low scores in the admission test are more likely to have trouble in achieving the necessary CFU by the end of the first year within the stipulated time scale. It is also worth mentioning that the results show that there is still space to improve the discriminatory capacity of the admission test in order to identify the best students to admit onto the course. This aspect could be improved by means of a longer observation period that would allow us to assess whether, and to what extent, the difference between the 'Regular' and 'TAR' students varies once the study subjects no longer reflect those of the admission test, allowing us to identify the possible areas in which changes are required in order to improve the admission test.
In conclusion, the results seem to highlight the importance of employing an evaluation instrument that permits students to evaluate their potential to undertake a certain degree course. Elimination of the admission test would allow many students to enrol on degree courses in medicine who, from the results of our study, would actually have a very low probability of achieving the final goal.

This would also expose students and medical schools, not only to enormous economic burdens, but also to significant human challenges. Moreover, medical schools would be obliged to remodel their courses in order to accommodate the increased number of students, without necessarily gaining adequate benefit from such an investment.

Contributorship

All the authors participated in the definition of the design of the study, analysis of the discussion and drafting of the final document. In particular the specific contributions are:

Giuseppe Migliaretti - design, analysis and discussion and drafting of the final document

Salvatore Bozzaro – discussion and drafting of the final document

Roberta Siliquini - discussion and drafting of the final document

Giuseppe Costa – design, analysis and discussion and drafting of the final document

Franco Cavallo - design, analysis and discussion and drafting of the final document

Competing interest: any competing interest

Funding: the work has not benefited from any kind of economic support

References

1. Ibrahim A. Al-Alwan Association between Scores in High School, Aptitude and Achievement Exams and Early Performance in Health Science College Saudi J Kidney Dis Transpl 2009;20(3):448-453
2. Nicholson, S. Commentary: The benefits of aptitude testing for selecting medical students. BMJ 2005;331:559-60.
3. Groves MA, Gordon J, Ryan G. Entry tests for graduate medical programs: Is it time to rethink? Med J Aust 2007;186(3):120-3.
4. Ting Donga, Kimberly A. Swygertb, Steven J. Durninga, Aaron Saguilc, William R. Gillilanda, David Cruessd, Kent J. DeZeea, Jeffrey LaRochellea & Anthony R. Artino Jr.a Validity Evidence for Medical School OSCEs: Associations With USMLE® Step Assessments. Teaching and Learning in Medicine, 2014. 26(4), 379–386
5. Wilkinson D, Zhang J, Byrne GJ, et al. Medical school selection criteria and the prediction of academic performance Evidence leading to change in policy and practice at the University of Queensland. Med J Aust 2008; 188:349-54.
6. Parry J, Mathers J, Stevens A, et al. Admissions processes for five year medical courses at English schools: Review. BMJ 2006;332:1005-9.
7. Albanese MA, Snow MH, Skochelack SE, Huggett KH, Farrell PM. Assessing personnel qualities in medical school admissions. Acad Med 2003;78:313-21.
8. Admission Requirements of Canadian Faculties of Medicine. www.afmc.ca/docs/2005AdBk.pdf (accessed 19 Jan 2006).
9. Coebergh J. Dutch medical schools abandon selection for lottery system for places. Student BMJ 2003;11:138.
10. Parry J, Mathers J, Stevens A, et al. Admissions processes for five year medical courses at English schools: Review. BMJ 2006;332:1005-9.
11. Hodges B, Regehr, G, Hanson, M, McNaughton N. Validation of an objective structured clinical examination in psychiatry. Academic Medicine 1998;73:910–2.
12. Park RS, Chibnall JT, Blaskiewicz RJ, Furman GE, Powell JK, Mohr CJ. Construct validity of an objective structured clinical examination (OSCE) in psychiatry: Associations with the clinical skills examination and other indicators. Academic Psychiatry 2004;28:122–8.
13. Walters K, Osborn D, Raven P. The development, validity and reliability of a multimodality objective structured clinical examination in psychiatry. Medical Education 2005;39:292–8.
14. Raj N, Badcock LJ, Brown GA, Deighton CM, O'Reilly SC. Design and validation of 2 objective structured clinical examination stations to assess core undergraduate examination skills of the hand and knee. The Journal of Rheumatology 2007;34:421–4.
15. Delavar M, Salmalian H, Faramarzi M, Pasha H, Bakhtiari A, Nikpour M, et al. Using the objective structured clinical examinations in undergraduate midwifery students. Journal of Medicine and Life 2013;6:76–9.
16. Berg K, Winward M, Clauser BE, Veloski JA, Berg D, Dillon GF, et al. The relationship between performance on a medical school's clinical skills assessment and USMLE Step 2 CS. Academic Medicine 2008;83: S37–40.
17. Bredo, E. Reconstructing educational psychology: Situated cognition and Deweyian pragmatism. Educational Psychology 1994;29:23–35.
18. Wilson BG, Myers KM. Situated cognition in theoretical and practical context. In D Jonassen & S Land (Eds.), Theoretical foundations of learning environments (pp. 57–88). Mahwah, NJ: Erlbaum, 2000.
19. Young MF. Instructional design for situated learning. Educational Technology Research and Development 1993;41:43–53.

Table 1. Number of students in each student category ('Regular' vs. 'TAR') attending the two University of Turin colleges

	REGULAR	TAR	TOTAL
Molinette	402	124	526
San Luigi	129	28	157
TOTAL	531	152	683

Table 1.1 Average score achieved in the admission test per student category (Total score and score per discipline)

		Biology	Chemistry	General culture	Physics & Mathematics	Logic	TOTAL
REGULAR	Mean	8.9	6.2	0.7	6.0	19.5	41.2
	S.D.	3.2	2.9	1.1	3.1	4.8	6.2
	Median	8.9	6.3	0.0	5.6	19.8	39.9
	N	445	445	445	445	445	445
TAR	Mean	5.1	2.8	0.3	2.3	13.3	23.9
	S.D.	3.6	2.7	1.0	2.4	4.3	6.7
	Median	5.5	2.5	0.0	1.8	13.4	25.0
	N	152	152	152	152	152	152
	Difference	3.8	3.4	0.4	3.7	6.2	17.3
95% CI	lower	3.2	2.9	0.1	3.1	5.3	16.2
	upper	4.4	3.8	0.6	4.2	7.1	18.5
	p-value	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*

Table 2. Secondary school final grade, CFU and 1st-year exam grade per student category and per college

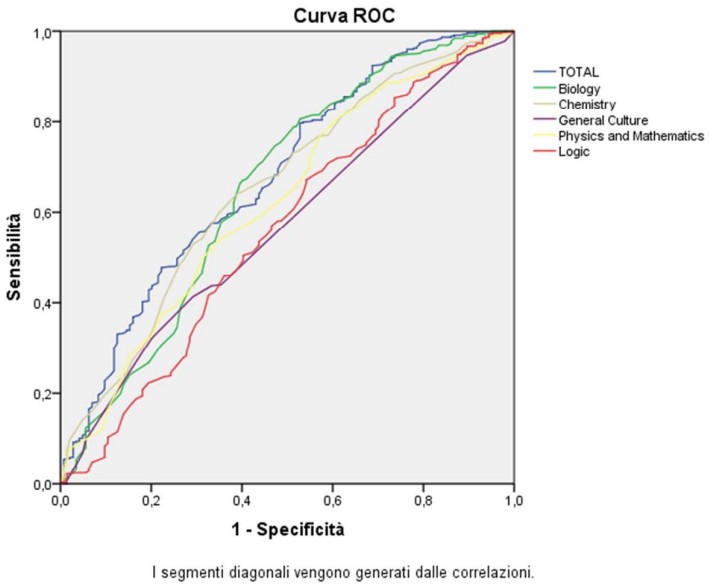
		TOTAL			MOLINETTE			SAN LUIGI		
		Secondary school final grade	CFU acquired	1st-year exam grade	Secondary school final grade	CFU acquired	1st-year exam grade	Secondary school final grade	CFU acquired	1st-year exam grade
REGULAR	Mean	88.4	36.7	26.5	88.5	37.3	26.6	87.9	35.1	26.3
	S.D.	10.2	16.1	2.1	10.3	17.2	2.0	9.8	12.3	2.4
	Median	89.0	43.0	26.8	90.0	44.0	26.9	89.0	37.0	26.5
	N	529	531	492	401	402	368	128	129	124
TAR	Mean	80.1	27.9	24.7	80.7	28.2	24.9	77.5	26.6	23.8
	S.D.	10.5	19.3	2.3	10.4	20.1	2.3	11.0	15.1	2.2
	Median	79.0	29.0	24.8	80.0	30.0	25.0	77.5	26.5	24.0
	N	152	152	123	124	124	100	28	28	23
	Difference	8.3	8.8	1.8	7.8	9.1	1.8	10.4	8.4	2.5
95% CI	lower	6.7	6.3	1.5	5.7	5.5	1.3	6.2	3.2	1.4
	upper	10.2	11.9	2.3	9.9	12.3	2.2	14.5	13.7	3.6
	p-value	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*0.002*	*<0.001*

Table 3. Predictive capacity of the admission test on the number of CFU acquired

	Gold Standard reference (GS)		Gold Standard reference (GS2)	
	P	N	P2	N2
REGULAR	419	112	217	314
TAR	89	63	30	122
TOTAL	508	175	247	436

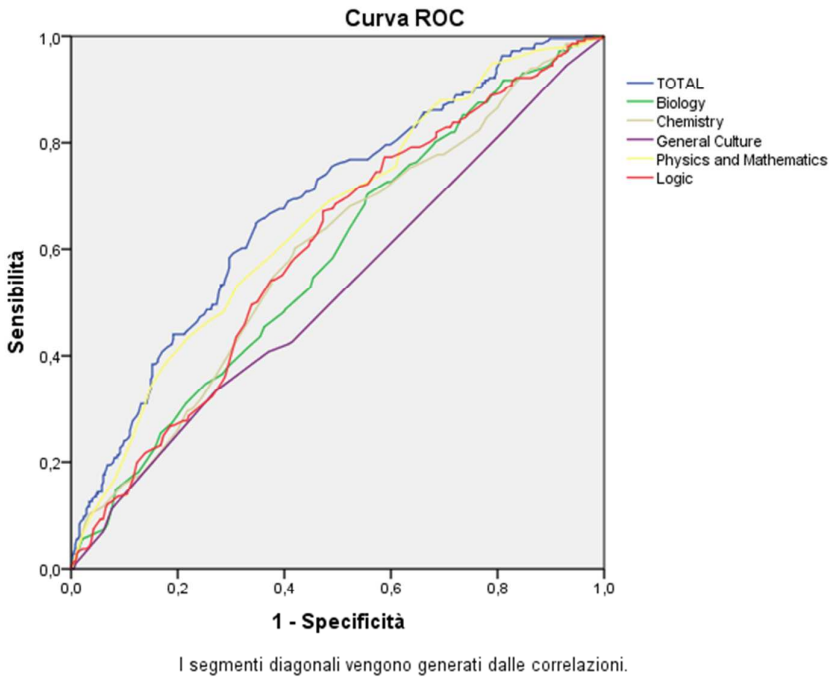
Se	0.82		0.88	
95%CI	0.79	0.86	0.83	0.91
Sp	0.36		0.28	
95%CI	0.29	0.43	0.24	0.32
PPV	0,79		0,41	
95%CI	0,74	0,82	0,37	0,45
NPV	0,41		0,8	
95%CI	0,34	0,59	0,73	0,86
LH+	1,29		1,22	
95%CI	1,17	1,41	1,15	1,29
LH-	0,49		0,43	
95%CI	0,37	0,60	0,36	0,61

Figure 1. ROC curve (Gold Standard reference: GS) –TOTAL and Subject Area Scores



	AUC	Standard Error	p-value	95% CI	
				lower	upper
TOTALE	,678	,026	,000	,626	,730
Biology	,653	,028	,000	,598	,709
Chemistry	,654	,027	,000	,602	,706
General Culture	,569	,027	,012	,517	,622
Physics and Mathematics	,622	,027	,000	,569	,676
Logic	,561	,029	,028	,504	,617

Figure 2. ROC curve (Gold Standard reference: GS2) –TOTAL and Subject Area Scores



Area under the curve

	AUC	Standard Error	p-value	Asymptotic 95% confidence interval	
				Lower	Upper
TOTALE	,681	,022	,000	,638	,725
Biology	,583	,024	,001	,536	,630
Chemistry	,592	,024	,000	,545	,639
General Culture	,523	,025	,360	,474	,571
Physics and Mathematics	,653	,023	,000	,608	,698
Logic	,601	,024	,000	,555	,648

Equator Network Research checklists – Diagnostic/Prognostic studies

TITLE OR ABSTRACT

1 Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC) ABSTRACT - OK. pag.2

2 Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts) OK – pag 2

INTRODUCTION

3 Scientific and clinical background, including the intended use and clinical role of the index test OK. Pag. 3

4 Study objectives and hypotheses METHODS Study design OK. Pag. 4

5 Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study) Participants OK. Pag. 4

6 Eligibility criteria OK. Pag. 4

7 On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry) OK. Pag. 4

8 Where and when potentially eligible participants were identified (setting, location and dates) OK. Pag. 4

9 Whether participants formed a consecutive, random or convenience series Test methods OK. Pag. 4

10a Index test, in sufficient detail to allow replication OK. Pag. 4

10b Reference standard, in sufficient detail to allow replication OK. Pag. 4

11 Rationale for choosing the reference standard (if alternatives exist) OK. Pag. 4

12a Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory OK. Pag. 4

12b Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory OK. Pag. 4

13a Whether clinical information and reference standard results were available to the performers/readers of the index test OK. Pag. 4

13b Whether clinical information and index test results were available to the assessors of the reference standard Analysis OK. Pag. 4

14 Methods for estimating or comparing measures of diagnostic accuracy OK. Pag. 4

15 How indeterminate index test or reference standard results were handled OK. Pag. 5

16 How missing data on the index test and reference standard were handled OK. Pag. 5

17 Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory OK. Pag. 4 - 5

18 Intended sample size and how it was determined RESULTS Participants OK. Pag. 4

19 Flow of participants, using a diagram All recruited students were included in the study

20 Baseline demographic and clinical characteristics of participants OK. Pag. 4

21a Distribution of severity of disease in those with the target condition. Not applicable

21b Distribution of alternative diagnoses in those without the target condition. Ok Pag 4

22 Time interval and any clinical interventions between index test and reference standard Test results Not applicable

23 Cross tabulation of the index test results (or their distribution) by the results of the reference standard OK. Pag.5-6 AND Table 3

24 Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals) OK. Pag.5-6 Table 3

25 Any adverse events from performing the index test or the reference standard DISCUSSION Not applicable

26 Study limitations, including sources of potential bias, statistical uncertainty, and generalisability OK pag. 2 AND Pag. 6-7

27 Implications for practice, including the intended use and clinical role of the index test OTHER INFORMATION OK. Pag 6-7

28 Registration number and name of registry Not applicable

29 Where the full study protocol can be accessed Not applicable

30 Sources of funding and other support; role of funders Study not economically supported – Pag. 7

BMJ Open

ADMISSION TEST FOR THE DEGREE COURSE IN MEDICINE AND ACADEMIC PERFORMANCE: AN ITALIAN EXPERIENCE FROM THE UNIVERSITY OF TURIN

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017417.R1
Article Type:	Research
Date Submitted by the Author:	21-Aug-2017
Complete List of Authors:	Migliaretti, Giuseppe; University of Turin, . Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin Bozzaro, Salvatore; University of Turin, Dept of Clinical and Biological Sciences, School of Medicine of University of Turin Siliquini, Roberta; University of Turin, Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin Costa, Giuseppe; University of Turin, Dept of Clinical and Biological Sciences, School of Medicine of University of Turin Cavallo, Franco; University of Turin, Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Epidemiology, Research methods
Keywords:	Admission test, Degree course in Medicine, academic performances, ROC curve, diagnostic test validity, University of Turin

ADMISSION TEST FOR THE DEGREE COURSE IN MEDICINE AND ACADEMIC PERFORMANCE: AN ITALIAN EXPERIENCE FROM THE UNIVERSITY OF TURIN

Migliaretti Giuseppe[1], Bozzaro Salvatore[2], Siliquini Roberta[1], Costa Giuseppe[2], Cavallo Franco[1]

1. Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin
2. Dept of Clinical and Biological Sciences, School of Medicine of University of Turin

Corresponding author

Giuseppe Migliaretti, Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin

Str. Santena 5bis, 10100 Turin (Italy)

Mail: giuseppe.migliaretti@unito.it

Phone: +39 11 6705808

Fax: +39 11 2365808

Co-Authors

Salvatore Bozzaro-Dept Clinical and Biological Sciences, School of Medicine of University of Turin (ITALY)

Roberta Siliquini-Dept Public Health and Pediatric Sciences, School of Medicine of University of Turin (ITALY)

Giuseppe Costa-Dept Clinical and Biological Sciences, School of Medicine of University of Turin (ITALY)

Franco Cavallo-Dept Public Health and Pediatric Sciences, School of Medicine of University of Turin (ITALY)

Keywords: Admission test, Degree course in Medicine, academic performances, ROC curve, Diagnostic test validity, University of Turin

Additional data: No additional data are available

Ethical aspects: This work has respected all ethical requirements and has not received any kind of economic support

WORD COUNT: 3970

ABSTRACT

Objectives

The usefulness of university admission tests to medical Schools has been discussed in recent years. In the academic year 2014/2015, several students who failed the admission test in Italy, appealed to the regional administrative court ('Tribunale Amministrativo Regionale' – TAR) requesting to be included, despite their test results. All won their appeals and were admitted to their respective courses. The existence of this population of students generated a control group, in order to evaluate the predictive capacity of the admission test. The aim of the present work is to discuss the ability of university admission tests to predict subsequent academic success.

Setting and Participants The present study considers the 683 students who enrolled onto the first year of the degree course in medicine in the academic year 2014/2015 at the University of Turin (Molinette and San Luigi Gonzaga colleges). The students were separated into two categories: those who passed the admission test [n1=531] and those who did not pass the admission test but won their appeal in the TAR [n2=152].

Outcomes

The validity of the admission test was analysed using Specificity, Sensitivity, Likelihood, ROC curves, Area under the ROC curve and relative 95% Confidence Intervals.

Results

The results show that the admission test appears to be a good tool for predicting the academic performances in the first year of course (AUC= 0.70, 95%CI from 0.64 to 0.76). Moreover, some subject areas seem to have a greater discriminating capacity than others. In general, students who obtain a high score in scientific questions are advantaged in reaching the required standards during the first year [LH+ = 1.22 (95% CI from 1.14 to 1.25)].

Conclusions

Based on consistent statistical approach, our study seems to confirm the ability of the admission test, to predict the academic success in the first year of the School of Medicine.

Strengths and limitations of this study

- This is the first study based on consistent statistical approach conducted in an Italian University to evaluate the reliability of the admission test
- The presence of a control group (TAR) improves the reliability of the results
- Comparability of the groups: the registration year, lessons, programmes, teachers and classrooms were the same for both Regular and TAR students.
- Further investigations will be required to analyse the two cohorts over a longer period of time
- The predictive capacity of the test was here only studied in relation to the academic performance on the basic courses; a relationship with clinical skills was not considered in this work

Bullet points

- The study confirms the ability of the admission test for the degree course in medicine to predict the subsequent academic performances in the first year of course
- A comparison was made between the academic performance of students who passed the admission test (*Regular)* with that of those who did not pass the admission test but who won their appeal in the 'Tribunale Amministrativo Regionale: TAR' (*TAR*)
- Statistical analyses based on ROC curves, Sensibility (Se) and Specificity (Sp) of the admission test, Positive Likelihood ratio (LH+), Negative Likelihood ratio (LH-) and relative 95% Confidence Intervals were computed
- The results show that the admission test appears to be a good tool for predicting which students will be able to achieve academic success in the first year.
- In conclusion, the results highlight the importance of an evaluation tool that allows students to assess their capacity to undertake a certain degree course.
- An assessment of the similarities between the Italian admission test and those of other countries, regarding question type and exam duration, will allow the results of this study to be extended to a more international context.

Introduction

In recent years, a greater demand for higher qualification created both economical and technical problems. Indeed, the number of classrooms, labs, infrastructures, technical staff and teachers must be sufficient to manage large numbers of students without compromising teaching quality as stressed in [1]. Moreover, the youth unemployment raises important questions about the number and the quality of graduates with respect to job opportunities, suggesting the need for more accurate selection procedures.

The usefulness of the admission tests, in particular for degree courses in Medicine, has been widely discussed in recent years, both in Europe [1–3] and in other continents [4]. The oldest admission tests come from USA (e.g. the Moss test in 1928), although different versions have been created over the years; the most recent test, called MCAT, was created in 2007 and is now used by almost all the colleges in North America [2]. As concerns European countries, no standards have been formulated to date. The European Union has only provided some general advice regarding the quality of education that aims toward progressive standardization. Each Country is thus allowed to take personalized actions: the English (called UKCAT) [3] and Austrian [1] tests were created in 2006, the Irish (called HPAT) test [2] was formulated in 2009, while France has never introduced an admission test, preferring to simply bar students who do not make the grade at the end of their first year from progressing in their course [5]. Other approaches have also been considered, including a totally open access to courses (e.g. Austria until 2002) [1] and random selection (e.g. Holland until 1999) [6]. In Italy, the admission test was proposed in 1987 by Zecchino (Minister of the Public Education) and finally introduced in 1999 as a law (264/99). The required skills, and how to test them, change from country to country (see Table 1).

Table 1: Comparison of admission tests from different countries.

	Italian admission test	**HPAT**	**UKCAT**	**UMAT**	**MCAT**
Country	Italy	Ireland	England	Australia	USA
Type of test	MCQ	MCQ	MCQ	MCQ	MCQ
Subsection of test	General culture Logic Biology Chemistry Mathematics and physics	Logical reasoning & problem solving Interpersonal understanding Nonverbal reasoning	Verbal reasoning Quantitative reasoning Abstract reasoning Decision analysis	Logical reasoning & problem solving Understanding people Nonverbal reasoning	Logical reasoning & problem solving
Duration	1h 40 min	2h 30 min	2h	2h 45 min	4h 30 min
Reference	[7]	[2]	[3]	[2]	[4]

A good review of the test types can be found in [4]. In brief, they can be divided in to cognitive, non-cognitive, written tests and interviews. Moreover, in some cases (USA, UK, Australia, Canada, New Zealand, and South Africa), considerable importance is given to ethnic minorities and disadvantaged groups in order to facilitate their admittance. The general trend is to use written cognitive tests, while almost all avoid interviews [4] because they are less predictive and more time- and money-consuming. In Italy, the test is administered nation-wide for the public universities, prepared by the Ministry of Education, Universities and Research (*Ministero dell'Istruzione, dell'Università e della Ricerca MIUR*). The test is written and composed of 60 multiple-choice questions (*MCQ*) to be answered within a limited time [7] (see Table 1 for details). Moreover, the access of disadvantaged groups and non-Italian people is guaranteed by the reservation of places for these categories.

In Italy, a debate is open about the present situation and the possibility to improve the admission test. Investigations into how the results of the admission tests can predict academic success are thus required.

Indeed, many European studies highlight a tight relationship between admission test results and academic performance [3,8], personal skills [2] or motivations [1]. However, other authors (e.g. [9]) have reported discordant results about the same tests. Unfortunately, the large majority of these studies have looked at situations outside of Italy [6,10–12]. Literature on the Italian admission test results is very scarce and is not specific to Medicine (see [13] about Nurses and [14] about Veterinary). The present statistical study (although regarding a single University) therefore makes an important contribution to the discussion on the usefulness of the admission test, both for Italian and non-Italian readers.

Another consideration that should be made regards the inner bias of the majority of past studies: i.e., the lack of adequate control groups. Indeed, only the results of students who had passed the admission test can be considered, and comparisons can only be made using previous cohorts, i.e. the students who entered the University without being tested [1]. However, in this case no information about the scores of the comparison group is present.

In Italy, in the academic year 2014/2015, many students, failing the admission test, appealed to the regional administrative court (Tribunale Amministrativo Regionale: TAR) because of supposed irregularities that occurred during the examination. The court accepted the appeal of recursive students on the basis of a ‘supposed infringement of anonymity principle’ and granted them admittance onto their respective courses. This ‘extraordinary’ situation produces a ‘control group’ of students (with lower scores) useful for evaluating the predictability of the admission test.

In the year 2016 a working group with the aim of assessing the predictive effectiveness of the admission test, was organised by the Italian Institution composed of Presidents of Italian Degree course in Medicine (named '*Permanent Conference of Presidents of Degree course in Medicine*' which allows homogeneity and coordination of the Schools of Medicine). Preliminary results of the working group's activities which will be published in the Conference's Journal (until now relative to just some Italian colleges), highlight the need for adjustments for the different colleges that are often not comparable with regard to study plans and the organization of courses and examinations.

Within this framework, the two constituent medical colleges of the University of Turin (Molinette and San Luigi Gonzaga) have long been developing this line of research on their students.

Objectives

The aim of the present work is to evaluate how the admission test can predict the subsequent academic success of medical students, taking advantage of the particular 2014/15 cohort of students from the University of Turin

Database

The study was approved by the degree course Council of the School in Medicine of the University of Turin (Molinette and San Luigi Gonzaga colleges) and by the students Committee. The analyses were performed on anonymized database (i.e. without sensitive data) provided by the medical schools.

The present study considers 683 students who enrolled onto the first year of the degree course in Medicine in the academic year 2014/2015 at the University of Turin (Molinette and San Luigi Gonzaga colleges). The students were divided into two categories: those who passed the admission test (*Regular*) [n1=531] and those who did not pass the admission test but won their appeal in the TAR (*TAR*) [n2=152].

All students were monitored until the end of the first year's last exam session (January 2015 – May 2016).

Methods

The following data are presented as means, standard deviations (SD), median and 95% confidence intervals for the two investigated student groups ('Regular' and 'TAR'): admission test score;

secondary school final grade; number of university exam credits (CFU) acquired; the average first-year exam grade.

Considering the number of CFU accumulated in the observation period as the principal end-point measure, students were classified into the following categories defining two different Reference Standard (RS):

RS1

- Students who acquired half, or more than half, of the required credits at the end of the first year (*P1*)
- Students who acquired less than half of the required credits at the end of the first year (*N1*)

RS2

- Students who acquired all the CFU required for the first year (*P2*)
- Students who did not acquire all the CFU in the first year (N2)

In order to evaluate the predictability of the admission test, the achievement of RS1 and RS2 (independent of the student category Regular and TAR) was evaluated using ROC curves [15–17]. Indeed, the ROC curve illustrates the ability of a test to discriminate the true positive (Sensitivity) from the false positive (1-Specificity) cases. If the test is strongly predictive, the curve grows rapidly. This shape should produce a large area under the ROC curve (AUC) reported with relative 95% Confidence Intervals (95%CI). An AUC value greater than 0.5 and close to 1 indicates a good level of predictability of the test.

In order to analyze the weight of each of the subject areas on the predictive capacity of the test, the ROC curve, the AUC and relative 95% CI were calculated for Total score and for each sub-area of the test.

Finally, we evaluated the 'goodness' of the cut-off score used at the University of Turin (33.9) for discriminating between admitted and non-admitted students. The analysis was based on Sensitivity (Se), Specificity (Sp), Positive Likelihood ratio (LH+), Negative Likelihood ratio (LH-) and relative 95% Confidence Intervals values [15,18,19].

Results

Table 2 shows the number, mean secondary school final grade, CFU, and mean 1st-year exam grade of the students enrolled in the first year according to group (Regular and TAR), whereas Table 3 presents the scores achieved in the admission tests (Total and per subject area). The results show that students comprising the Regular group obtained higher test scores and more CFU at the end of the first year than TAR students.

Table 2. Number of students, secondary school final grade, CFU and mean 1st-year exam grade in each student category ('Regular' vs. 'TAR') attending the two Universities of Turin

		TOTAL			MOLINETTE			SAN LUIGI		
		Secondary school final grade (max 100)	CFU acquired (max 51)	Mean 1st-year exam grade (max 30L)	Secondary school final grade (max 100)	CFU acquired (max 51)	Mean 1st-year exam grade (max 30L)	Secondary school final grade (max 100)	CFU acquired (max 51)	Mean 1st-year exam grade (max 30L)
REGULAR	Mean (SD)	88,2 (10,2)	36 (16,0)	26,5 (2,1)	88,5 (10,3)	37,3 (17,2)	26,6 (2,0)	87,9 (9,8)	35,1 (12,3)	26,3 (2,4)
	Median	89	40	26,7	90	44	26,9	89	37	26,5
	N	531	531	492	402	402	368	129	129	124
TAR	Mean (SD)	79,1 (10,6)	27,4 (18,9)	24,5 (2,3)	80,7 (10,4)	28,2 (20,1)	24,9 (2,3)	77,5 (11,0)	26,6 (15,1)	23,8 (2,2)
	Median	78	27	25	80	30	25	77,5	26,5	24
	N	152	152	123	124	124	100	28	28	23
Mean Difference (95CI%)		8,4 (from 6,7 to 10,2)	9,1 (from 6,3 to 11,9)	1,9 (from 1,5 to 2,3)	7,8 (from 5,7 to 9,9)	9,1 (from 5,5 to 12,3)	1,8 (from 1,3 to 2,2)	10,4 (from 6,2 to 14,5)	8,4 (from 3,2 to 13,7)	2,5 (from 1,4 to 3,6)

Table 3 Average score achieved in the admission test per student category (Total and per Subject area)

		Biology	**Chemistry**	**General culture**	**Physics &** Mathematics	**Logic**	**TOTAL**
REGULAR	Mean (SD)	8,9 (3,2)	6,2 (2,9)	0,7 (1,1)	6 (3,1)	19,5 (4,8)	41,2 (6,2)
	Median	8,9	6,3	0	5,6	19,8	39,9
	N	531	531	531	531	531	531
TAR	Mean (SD)	5,1 (3,6)	2,8 (2,7)	0,3 (1,0)	2,3 (2,4)	13,3 (4,3)	23,9 (6,7)
	Median	5,5	2,5	0	1,8	13,4	25
	N	152	152	152	152	152	152
Mean Difference (95CI%)		3,8 (from 3,2 to 4,4)	3,4 (from 2,9 to 3,8)	0,4 (from 0,1 to 0,6)	3,7 (from 3,1 to 4,2)	6,2 (from 5,3 to 7,1)	17,3 (from 16,2 to 18,5)

Based on the ROC curves analysis, the admission test appears to be a good tool (AUC= 0.67, 95% CI from 0.63 to 0.71) for predicting RS1 achievement (Table 4). The specific analysis performed for sub-areas of the test show that some of them seem to have a greater discriminating capacity than others: Biology (AUC=0.61, 95% CI from 0.57 to 0.66); Physics/Mathematics (AUC=0.63, 95% CI from 0.59 to 0.68); and Chemistry (AUC=0.65, 95% CI from 0.61 to 0.69) (Table 4).

One again, the admission test appears to be predictive of RS2 achievement (AUC= 0.70, 95% CI from 0.64 to 0.76) (Table 4). Also in this case some sub-areas have a greater discriminating capacity than other: Biology (AUC=0.62, 95% CI from 0.56 to 0.68), Physics/Mathematics (AUC=0.63, 95% CI from 0.57 to 0.69) and Chemistry (AUC=0.63, 95% CI from 0.57 to 0.69) (Table 4).

Table 4. Area under ROC curve (AUC) and relative 95% Confidence Interval for Total and per Subject area scores

	RS1 *AUC (95%CI)*	**RS2** *AUC (95%CI)*
TOTAL	0,67 (from 0,63 to 0,71)	0,70 (from 0,64 to 0,76)
Biology	0,61 (from 0,57 to 0,66)	0,62 (from 0,56 to 0,68)
Chemistry	0,65 (from 0,61 to 0,69)	0,63 (from 0,57 to 0,69)
General culture	0,55 (from 0,50 to 0,59)	0,57 (from 0,51 to 0,64)
Physics&Mathematics	0,63 (from 0,59 to 0,68)	0,63 (from 0,57 to 0,69)
Logic	0,54 (from 0,50 to 0,60)	0,59 (from 0,53 to 0,65)

Specific analysis of the ROC curve numerical values shows a value between 30 and 35 to be a good cut-off score to identify the 'admitted' students, which is in agreement with the institutional one used by the Turin colleges (33.9 in the year 2014/15). In order to understand the reliability of the test, the real classification used at the University of Turin, **i.e.** Regular and TAR students (based on cut-off score equal to 33.9) was evaluated in relation to Reference Standard RS1 and RS2 (Table 5).

For both RS1 and RS2, Sensitivity is high (Se = 0.86, 95% CI from 0.82 to 0.89 for RS1; Se = 0.91, 95% CI from 0.84 to 0.95 for RS2), whereas Specificity is low (Sp = 0.31, 95% CI from 0.26 to 0.36 for RS1; Sp = 0.25, 95% CI from 0.21 to 0.28 for RS2). Analogously, positive Likelihood ratios LH+ are both higher than 1 (1.23, 95% CI from 1.15 to 1.32 for RS1; 1.22, 95% CI from 1.14 to 1.25 for RS2), while negative Likelihood ratios LH- are both less than 1 (0.46, 95% CI from 0.22 to 0.54 for RS1; 0.36, 95% CI from 0.23 to 0.56 for RS2).

The high values of Se indicate that the majority of students who achieve the RSs are from the Regular group, while a low Sp indicates that a number of 'Regular' students are also among those who did not achieve the two RSs (see Table 5). This interpretation is also supported by the good value of LH+, which also points out that being 'Regular' is a 'protective factor' to achieving both RS1 and RS2.

For our study, however, it is essential to highlight the low value of LH-, which points out that relatively few TAR students achieved the RSs (in particular RS2). This indicates that a low score in the test also predicts a low probability of achieving the RSs.

Table 5. Predictive capacity of the admission test on the number of CFU acquired

	RS1		RS2	
	P1*	N1**	$P2^{+}$	$N2^{++}$
REGULAR	308	223	101	430
TAR	52	100	10	142
TOTAL	360	323	111	572

Se (95% CI)	0,86 (from 0,82 to 0,89)	0,91 (from 0,84 to 0,95)
Sp (95%CI)	0,31 (from 0,26 to 0,36)	0,25 (from 0,21 to 0,28)
LH+ (95% CI)	1,23 (from 1,15 to 1,32)	1,22 (from 1,14 to 1,29)
LH- (95% CI)	0,46 (from 0,22 to 0,54)	0,36 (from 0,23 to 0,56)

**P1: students who acquired half, or more than half, of the required credits at the end of the first year*

***N1: students who acquired less than half of the required credits at the end of the first year*

$^{+}$P2: students who acquired all the CFU required for the first year

$^{++}$N2: students who did not acquire all the CFU in the first year

Discussion

The admission of a significant number of students to the degree course in medicine at the University of Turin who did not pass the admission test in the year 2013/14 gave us, for the first time, the opportunity to compare the academic results of two student groups in the same context. A close correlation between academic success and test results was found; in particular, students who got a high score in scientific questions were more likely to achieve the requested standards during the first year. Although it could be improved, the admission test therefore seems to be a good tool for identifying the better students.

Other studies have analyzed the performance of students who passed an admission test versus open access students (see, for example, [1]), but the comparison could only be made between students of different years. In our case, registration year, lessons, programmes, teachers and classroom characteristics were the same for both Regular and TAR students. For the purposes of our study, it is important to point out that 80% of the TAR students achieved a score of 20 to 31 points in the test (the minimum score for the admission to Turin's School of Medicine was 33.9). This substantiates the TAR group as a good reference.

The Italian test is comparable to other tests used in international contexts, at least in relation to question type and exam duration (Table 1), and for this reason our results may also be of interest outside Italy.

It is important to highlight the similar results found in terms of usefulness and predictability. For example, [3] underlined a clear correlation between 'total science score' (our Biology, Chemistry and Physics/Mathematic scores) and 'education performance' (our RS2, although in [3] a longer observational period was considered).

Most of the previous studies [1–3] confirmed that admission tests are able to predict the academic results in the first year. Nevertheless, no generalized predictability is assessed, also because of the variety of the evaluation periods and reference standards. For example, [3] evaluated the overall career of the students, while [1,2] and our study considered the results of the first or second years. As far as the reference standards, [1] investigated the dropout rates (lower in the students passing the admission test), whereas [2] was interested in prediction criteria for clinical and communication skills.

The strengths of this study are the presence of a reliable control cohort (TAR) and the possibility of specific analysis per admission test subject area. Although several debates are ongoing in Italy regarding which specific subject areas are most useful for discriminating between potential medical students, our study shows that the results for questions on biology, chemistry and physics/mathematics in the current admission test present the best predictability.

Several limitations of our study should however be taken into consideration. First of all, we only considered students admitted to the course of Medicine in Turin and not a wider Italian cohort. Second, this work constitutes an initial explorative analysis, limited to the first year of the medical course.

It is worth pointing out that the admission test is predictive of academic success, but not necessarily the ability to practice as a doctor. This aspect was stressed by many authors [20–22], but no definitive conclusion has been reached. In general, however, a single admission test is unable to predict ability to practice as a physician.

Important information could be obtained by following our two cohorts over a longer period of time. The development of this study shall, in fact, monitor the two cohorts throughout the complete academic path, re-evaluating their results also with respect to subjects (e.g. clinically oriented courses) different from those considered in the admission test.

Conclusions

In conclusion, admission test scores are able to predict the academic success in the first year of the medical course; this is an advantage for both students and medical schools. Indeed, it discourages students who do not pass the test from enrolling in the course, driving them towards alternative courses, and saving them both time and money. With this selection procedure, universities are able to manage a lower number of more motivated students with higher probabilities of obtaining success. This allows a more efficient use of infrastructural and personnel resources. However, the discriminatory capacity of the admission test could be improved by replicating the analysis presented at the end of the fourth and sixth years, investigating the relationship between admission test results and clinical skills.

Contributorship

All authors contributed toward study design, data analysis and discussion and the drafting of the final manuscript. Specific contributions are as follows:

Giuseppe Migliaretti - design, analysis and discussion and drafting of the final document

Salvatore Bozzaro – discussion and drafting of the final document

Roberta Siliquini - discussion and drafting of the final document

Giuseppe Costa – design, analysis and discussion and drafting of the final document

Franco Cavallo - design, analysis and discussion and drafting of the final document

Aknowledgement

We wish to thank for their valued support:

Prof Anna Bossi – Dept of Clinical Sciences and Community, University of Milan

Prof Anita Tabacco – DISMA Dept of Mathematical Sciences "G. L. Lagrange", Polytechnic of Turin

Prof Caterina Guiot - Dept of Neurosciences, University of Turin

Dr Barbara Martin – Dept of Chemistry, University of Turin

Competing interest: no competing interest

Funding: the work has not received any kind of economic support

References

1. Reibnegger, G. *et al.* Dropout rates in medical students at one school before and after the installation of admission tests in Austria. *Acad. Med. J. Assoc. Am. Med. Coll.* **86,** 1040–1048 (2011).
2. Kelly, M. E. *et al.* To what extent does the Health Professions Admission Test-Ireland predict performance in early undergraduate tests of communication and clinical skills? An observational cohort study. *BMC Med. Educ.* **13,** 68 (2013).
3. Sartania, N., McClure, J. D., Sweeting, H. & Browitt, A. Predictive power of UKCAT and other pre-admission measures for performance in a medical school in Glasgow: a cohort study. *BMC Med. Educ.* **14,** 116 (2014).
4. Prideaux, D. *et al.* Assessment for selection for the health care professions and specialty training: consensus statement and recommendations from the Ottawa 2010 Conference. *Med. Teach.* **33,** 215–223 (2011).
5. Segouin, C. *et al.* Country report: medical education in France. *Med. Educ.* **41,** 295–301 (2007).
6. Coebergh, J. Dutch medical schools abandon selection for lottery system for places | Student BMJ. **11,** 138
7. Decreto Ministeriale del 29 novembre 2013 n. 986 - Atti Ministeriali MIUR. Available at: http://attiministeriali.miur.it/anno-2013/novembre/dm-29112013.aspx. (Accessed: 24th July 2017)
8. Al-Alwan, I. A. Association between scores in high school, aptitude and achievement exams and early performance in health science college. *Saudi J. Kidney Dis. Transplant. Off. Publ. Saudi Cent. Organ Transplant. Saudi Arab.* **20,** 448–453 (2009).
9. Yates, J. & James, D. The UK Clinical Aptitude Test and clinical course performance at Nottingham: a prospective cohort study. *BMC Med. Educ.* **13,** 32 (2013).
10. Parry, J. *et al.* Admissions processes for five year medical courses at English schools: review. *BMJ* **332,** 1005–1009 (2006).
11. Delavar, M. A. *et al.* Using the objective structured clinical examinations in undergraduate midwifery students. *J. Med. Life* **6,** 76–79 (2013).
12. Berg, K. *et al.* The relationship between performance on a medical school's clinical skills assessment and USMLE Step 2 CS. *Acad. Med. J. Assoc. Am. Med. Coll.* **83,** S37-40 (2008).

13. Lancia, L., Petrucci, C., Giorgi, F., Dante, A. & Cifone, M. G. Academic success or failure in nursing students: results of a retrospective observational study. *Nurse Educ. Today* **33,** 1501–1505 (2013).

14. Mannella, R. The Italian veterinary medicine admission test: analysis of student intake in the years 2007, 2008, and 2009, and of the test's relationship with students' academic careers. *J. Vet. Med. Educ.* **38,** 184–193 (2011).

15. Altman, D. *Practical statistics for medical research*. (Chapman & Hall).

16. Hanley, J. A. & McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143,** 29–36 (1982).

17. Zweig, M. H. & Campbell, G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.* **39,** 561–577 (1993).

18. Harper, R. & Reeves, B. Reporting of precision of estimates for diagnostic accuracy: a review. *BMJ* **318,** 1322–1323 (1999).

19. Nam, J. □M. Confidence Limits for the Ratio of Two Binomial Proportions Based on Likelihood Scores: Non□Iterative Method. *Biom. J.* **37,** 375–379 (1995).

20. Louridas, M., Szasz, P., de Montbrun, S., Harris, K. A. & Grantcharov, T. P. Can We Predict Technical Aptitude?: A Systematic Review. *Ann. Surg.* **263,** 673–691 (2016).

21. Kreiter, C. D. & Axelson, R. D. A perspective on medical school admission research and practice over the last 25 years. *Teach. Learn. Med.* **25 Suppl 1,** S50-56 (2013).

22. Patterson, F. *et al.* How effective are selection methods in medical education? A systematic review. *Med. Educ.* **50,** 36–60 (2016).

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract **OK (pag.1-2)**	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported **OK (pag.4-5 and pag.6 rows 1-6)**
Objectives	3	State specific objectives, including any prespecified hypotheses **OK (pag.6 rows 8-11)**
Methods		
Study design	4	Present key elements of study design early in the paper **OK (pag.6 rows 27-32 and pag.7 rows 1-24 and Ref. 15-19)**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **OK (pag.6 rows 19-31)**
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls **OK (pag.5 rows 26-31 and pag.6 rows 19-23)** *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable **OK (pag.6 rows 19-31)**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **OK (pag.6 rows 13-31 and pag.7 rows 1-24)**
Bias	9	Describe any efforts to address potential sources of bias **OK (pag.11 rows 16-27)**
Study size	10	Explain how the study size was arrived at **OK (pag.5 rows 26-31 and pag.6 rows 19-25)**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why **OK (pag.6 rows 28-31)**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding **OK (pag.7 rows 1-24)**
		(*b*) Describe any methods used to examine subgroups and interactions **not appl**
		(*c*) Explain how missing data were addressed **NO Missing data**
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Results		
Participants **OK (pag.5 rows 26-31 and pag.6 rows 19-23)**	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders **OK (Table 2 and 3)**
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure **OK (pag.6 rows 19-23 and Table 2 and Table 3)**
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results **OK (pag.7 rows 26-31 and pag.8-9)**	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses **OK**	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives **OK (pag.10 rows 19-25)**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias **OK (pag.11 rows16-28)**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence **OK (pag.10 rows 29-33)**
Generalisability	21	Discuss the generalisability (external validity) of the study results **OK (pag.10 rows 34-36 and pag.11 rows1-18)**
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based **Not appl**

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

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

BMJ Open

Is the admission test for the course in Medicine a good predictor of academic performance? A 'case-control' experience at the School of Medicine of Turin

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017417.R2
Article Type:	Research
Date Submitted by the Author:	29-Sep-2017
Complete List of Authors:	Migliaretti, Giuseppe; University of Turin, . Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin Bozzaro, Salvatore; University of Turin, Dept of Clinical and Biological Sciences, School of Medicine of University of Turin Siliquini, Roberta; University of Turin, Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin Costa, Giuseppe; University of Turin, Dept of Clinical and Biological Sciences, School of Medicine of University of Turin Cavallo, Franco; University of Turin, Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Epidemiology, Research methods
Keywords:	Admission test, Degree course in Medicine, academic performances, ROC curve, diagnostic test validity, University of Turin

Is the admission test for the course in Medicine a good predictor of academic performance? A 'case-control' experience at the School of Medicine of Turin

Migliaretti Giuseppe[1], Bozzaro Salvatore[2], Siliquini Roberta[1], Costa Giuseppe[2], Cavallo Franco[1]

1. Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin
2. Dept of Clinical and Biological Sciences, School of Medicine of University of Turin

Corresponding author

Giuseppe Migliaretti, Dept of Public Health and Pediatric Sciences, School of Medicine of University of Turin

Str. Santena 5bis, 10100 Turin (Italy)

Mail: giuseppe.migliaretti@unito.it

Phone: +39 11 6705808

Fax: +39 11 2365808

Co-Authors

Salvatore Bozzaro-Dept Clinical and Biological Sciences, School of Medicine of University of Turin (ITALY)

Roberta Siliquini-Dept Public Health and Pediatric Sciences, School of Medicine of University of Turin (ITALY)

Giuseppe Costa-Dept Clinical and Biological Sciences, School of Medicine of University of Turin (ITALY)

Franco Cavallo-Dept Public Health and Pediatric Sciences, School of Medicine of University of Turin (ITALY)

Keywords

Admission test, Degree course in Medicine, Academic performances, ROC curve, Diagnostic test validity, University of Turin

Data Sharing Statement

No additional data are available

Ethical aspects

This work has respected all ethical requirements and has not received any kind of economic support

WORD COUNT: 2804

ABSTRACT

Objectives

The usefulness of university admission tests to medical schools has been discussed in recent years. In the academic year 2014-2015 in Italy, several students who failed the admission test, appealed to the regional administrative court ('Tribunale Amministrativo Regionale' – TAR) requesting to be included, despite their test results and all were admitted to their respective courses. The existence of this population of students generated a control group, in order to evaluate the predictive capacity of the admission test. The aim of the present work is to discuss the ability of university admission tests to predict subsequent academic success.

Setting and Participants The study considers the 683 students who enrolled onto the first year of the degree course in medicine in the academic year 2014-2015 at the University of Turin (Molinette and San Luigi Gonzaga colleges). The students were separated into two categories: those who passed the admission test [n1=531] and those who did not pass the admission test but won their appeal in the TAR [n2=152].

Outcomes

The validity of the admission test was analysed using Specificity, Sensitivity, Likelihood, ROC curves, Area under the ROC curve and relative 95% Confidence Intervals.

Results

The results show that the admission test appears to be a good tool for predicting the academic performances in the first year of course (AUC= 0.70, 95%CI from 0.64 to 0.76). Moreover, some subject areas seem to have a greater discriminating capacity than others. In general, students who obtain a high score in scientific questions are more likely to obtain the required standards during the first year [LH+ = 1.22 (95% CI from 1.14 to 1.25)].

Conclusions

Based on consistent statistical approach, our study seems to confirm the ability of the admission test, to predict the academic success in the first year of the School of Medicine.

Strengths and limitations of this study

- This is the first study based on a consistent statistical approach and conducted in an Italian University to evaluate the reliability of the Italian university admission test
- The presence of the TAR control group improves the reliability of the results
- Comparability of the groups: the registration year, lessons, programmes, teachers and classrooms were the same for both Regular and TAR students rendering the comparison of the two groups highly valid.
- Further investigations will be required to analyse the two cohorts over a longer period of time
- The predictive capacity of the test was only studied here in relation to academic performance in the basic courses; a relationship with clinical skills was not considered in this work

1. INTRODUCTION

In recent years, the growing demand for higher qualifications has created both economical and technical problems. Indeed, as discussed by Reibnegger and colleagues[1], the number of classrooms, labs, infrastructures, technical staff and teachers has had to increase contemporaneously in order to manage large numbers of students without compromising teaching quality. Moreover, the level of youth unemployment has raised important questions about the number and quality of graduates with respect to job opportunities, suggesting the need for more stringent selection procedures.

The usefulness of university admission tests, in particular for degree courses in Medicine, has been widely discussed in recent years, both in Europe [1–3] and in other continents [4]. The USA was the first to use admission tests in student selection procedures (e.g. the Moss test in 1928) and different versions have been created over the years. The most recent test, denominated MCAT, was created in 2007 and is now used by almost all the colleges in North America [2]. As concerns European countries, no standards have been formulated to date. The European Union has provided general advice only regarding the quality of education that point towards progressive standardization. Each Country is thus allowed to take personalized actions: the English (UKCAT [3]) and Austrian[1] tests were created in 2006, the Irish (HPAT) test [2] was formulated in 2009, while France has never introduced an admission test, preferring the strategy of simply barring students who do not make the grade at the end of their first year from progressing in their course[5]. Other approaches have also been considered, including a totally open access to courses (e.g., as applied in Austria until 2002) [1] and random selection (e.g., in Holland until 1999) [6]. In Italy, the admission test was proposed in 1987 by Zecchino (the Minister for Public Education) and finally introduced in 1999 as a law (264/99). The required skills, and how to test them, change from country to country (see Table 1).

Table 1: Comparison of admission tests in different countries.

	Italian admission test	**HPAT**	**UKCAT**	**UMAT**	**MCAT**
Country	Italy	Ireland	England	Australia	USA
Type of test	MCQ	MCQ	MCQ	MCQ	MCQ
Subsection of test	General culture Logic Biology Chemistry Mathematics and physics	Logical reasoning & problem solving Interpersonal understanding Nonverbal reasoning	Verbal reasoning Quantitative reasoning Abstract reasoning Decision analysis	Logical reasoning & problem solving Understanding people Nonverbal reasoning	Logical reasoning & problem solving

Duration	1h 40 min	2h 30 min	2h	2h 45 min	4h 30 min
Reference	Decreto Ministeriale n. 986 [7]	Kelly M. E. *et al.* [2]	Sartania N. *et al* [3]	Kelly M. E. *et al.*[2]	Prideaux, D. *et al.*[4]

A good review of the test types can be found in [4]. In brief, the tests can be divided into cognitive, non-cognitive, and written tests and interviews. Moreover, in some cases (the USA, the UK, Australia, Canada, New Zealand and South Africa), considerable importance is given to ethnic minorities and disadvantaged groups in order to facilitate their admittance. The general trend is to use written cognitive tests, while almost all avoid interviews [4] because they are less predictive and more time- and money-consuming. In Italy, the test is administered nation-wide for the public universities, prepared by the Ministry for Education, Universities and Research (*Ministero dell'Istruzione, dell'Università e della Ricerca MIUR*). The test is written and composed of 60 multiple-choice questions (*MCQ*) to be answered within a limited time [7] (see Table 1 for details). Moreover, the access of disadvantaged groups and non-Italian people is guaranteed by the reservation of places for these categories.

In Italy, a debate is open about the present situation and the possibility to improve the admission test. Investigations into how the results of the admission tests can predict academic success are thus required. Indeed, many European studies highlight a tight relationship between admission test results and academic performance [3,8], personal skills [2] and level of motivation[1]. However, other authors (e.g. [9]) have reported discordant results regarding the same tests. The large majority of these studies have concerned situations outside of Italy [6,10–12]. Thus, literature on the admission test results in Italy is very scarce and not specific to Medicine (see [13] about Nurses and [14] about Veterinary). The present statistical study (although regarding a single University) therefore makes an important contribution to the discussion on the usefulness of admission tests, both for Italian and non-Italian readers.

Another consideration that should be made regards the selection bias of the majority of past studies: i.e., the lack of adequate control groups. Indeed, only the results of students who had passed the admission test could be considered (as those failing the test were not admitted to university), and comparisons could only be made against previous cohorts (i.e., the students who entered the University without being tested)[1]; however, in this case no information about the scores of the comparison group is present.

In Italy, in the academic year 2014-2015, many of the students who failed the admission test appealed to the regional administrative court (Tribunale Amministrativo Regionale: TAR) on the account of supposed irregularities that occurred during the examination. The court accepted the appeal of recursive students on the basis of a 'supposed infringement of anonymity principle' and granted them admittance onto their respective courses. This 'extraordinary' situation generated a

'control group' of students (with lower test scores) useful for evaluating the predictability of the admission test.

Objectives

The aim of the present work was to evaluate whether the university admission test for the degree course in medicine can predict subsequent academic success in medical students, taking advantage of the particular 2014-2015 cohort of students from the University of Turin

2. METHODS

Setting

In 2016 a working group with the aim of assessing the predictive effectiveness of the admission test, was organised by the Italian Institution composed of Presidents of Italian Degree course in Medicine (named '*Permanent Conference of Presidents of Degree course in Medicine*' which allows homogeneity and coordination of the Schools of Medicine). Preliminary results of the group's activities (relative to just some Italian colleges) and which will be published in the Permanent Conference's Journal, highlight the need for changes to be made to different colleges, which do not always propose comparable curriculums and whose courses and examinations are organized differently. Within this framework, the two constituent medical colleges of the University of Turin (Molinette and San Luigi Gonzaga) have long been developing this line of research on their students.

Database

The study was approved by the Degree Course Council for the School in Medicine of the University of Turin (Molinette and San Luigi Gonzaga colleges) and by the Students' Committee. Approval by an ethical board is not explicitly required in Italy when the analysis of retrospective data is carried out, especially when data do not deal with disease conditions or use of pharmaceutical products. In order to meet the requirements of the Helsinki Declaration, the analyses were performed on anonymized database (without sensitive data) provided directly by the Medical Schools.

The present study considers 683 students who enrolled onto the first year of the degree course in Medicine in the academic year 2014-2015 at the University of Turin (Molinette and San Luigi

Gonzaga colleges). The students were divided into two categories: those who passed the admission test (*Regular*) [n1=531] and those who did not pass the admission test but won their appeal in the TAR (*TAR*) [n2=152].

All students were monitored until the end of the first year's last exam session (January 2015 – May 2016).

Statistical methods

The following data are presented as means, standard deviations (SD), median and 95% confidence intervals for the two investigated student groups ('Regular' and 'TAR'): admission test score; secondary school final grade; number of university exam credits (CFU) acquired; the average first-year exam grade.

Considering the number of CFU accumulated in the observation period as the principal end-point measure, students were classified into the following categories defining two different Reference Standard (RS):

RS1

- Students who acquired half, or more than half, of the required credits at the end of the first year (*P1*)
- Students who acquired less than half of the required credits at the end of the first year (*N1*)

RS2

- Students who acquired all the CFU required for the first year (*P2*)
- Students who did not acquire all the CFU in the first year (N2)

In order to evaluate the predictability of the admission test, the achievement of RS1 and RS2 (independent of the student category Regular and TAR) was evaluated using ROC curves [15–17]. Indeed, the ROC curve illustrates the ability of the admission test to discriminate true positive cases (Sensitivity) from false positive (1-Specificity) cases. If the test has high predictive capacity the curve grows rapidly; this shape should produce a large area under the ROC curve (AUC) reported with relative 95% Confidence Intervals (95%CI). An AUC value greater than 0.5 and close to 1 indicates a good level of predictability of the test.

In order to analyse the weight of each of the subject areas on the predictive capacity of the test, the ROC curve, the AUC and relative 95%CI were calculated for total score and for the individual sub-areas of the test.

Finally, we evaluated the 'goodness' of the cut-off score used at the University of Turin (33.9) for discriminating between admitted and non-admitted students. The analysis was based on Sensitivity

(Se), Specificity (Sp), Positive Likelihood ratios (LH+), Negative Likelihood ratios (LH-) and relative 95%CI values [15,18,19].

3. RESULTS

Table 2 shows the number, mean secondary school final grade, CFU, and mean first year exam grade of the students enrolled in the first year according to group (Regular and TAR), whereas Table 3 presents the scores achieved in the admission tests (Total and per subject area). The results show that students comprising the Regular group obtained higher test scores and more CFU at the end of the first year than TAR students.

Table 2. Number of students, secondary school final grade, CFU and mean first year exam grade in each student category ('Regular' vs. 'TAR') attending the two Universities of Turin

		TOTAL			MOLINETTE			SAN LUIGI		
		Secondary school final grade (max 100)	CFU acquired (max 51)	Mean first year exam grade (max 30L)	Secondary school final grade (max 100)	CFU acquired (max 51)	Mean first year exam grade (max 30L)	Secondary school final grade (max 100)	CFU acquired (max 51)	Mean first year exam grade (max 30L)
REGULAR	Mean (SD)	88.2 (10.2)	36 (16.0)	26.5 (2.1)	88.5 (10.3)	37.3 (17.2)	26.6 (2.0)	87.9 (9.8)	35.1 (12.3)	26.3 (2.4)
	Median	89	40	26.7	90	44	26.9	89	37	26.5
	N	531	531	492	402	402	368	129	129	124
TAR	Mean (SD)	79.1 (10.6)	27.4 (18.9)	24.5 (2.3)	80.7 (10.4)	28.2 (20.1)	24.9 (2.3)	77.5 (11.0)	26.6 (15.1)	23.8 (2.2)
	Median	78	27	25	80	30	25	77.5	26.5	24
	N	152	152	123	124	124	100	28	28	23
Mean Difference (95CI%)		8.4 (from 6.7 to 10.2)	9.1 (from 6.3 to 11.9)	1.9 (from 1.5 to 2.3)	7.8 (from 5.7 to 9.9)	9.1 (from 5.5 to 12.3)	1.8 (from 1.3 to 2.2)	10.4 (from 6.2 to 14.5)	8.4 (from 3.2 to 13.7)	2.5 (from 1.4 to 3.6)

Table 3 Average score achieved in the admission test per student category (total and per subject area)

		Biology	**Chemistry**	**General culture**	**Physics &** Mathematics	**Logic**	**TOTAL**
REGULAR	Mean (SD)	8.9 (3.2)	6.2 (2.9)	0.7 (1.1)	6 (3.1)	19.5 (4.8)	41.2 (6.2)
	Median	8.9	6.3	0	5.6	19.8	39.9
	N	531	531	531	531	531	531

TAR	Mean (SD)	5.1 (3.6)	2.8 (2.7)	0.3 (1.0)	2.3 (2.4)	13.3 (4.3)	23.9 (6.7)
	Median	5.5	2.5	0	1.8	13.4	25
	N	152	152	152	152	152	152
	Mean Difference (95CI%)	3.8 (from 3.2 to 4.4)	3.4 (from 2.9 to 3.8)	0.4 (from 0.1 to 0.6)	3.7 (from 3.1 to 4.2)	6.2 (from 5.3 to 7.1)	17.3 (from 16.2 to 18.5)

Based on the ROC curves analysis, the admission test appears to be good at predicting RS1 achievement (AUC= 0.67, 95% CI from 0.63 to 0.71; Table 4). The specific analysis performed for different sub-areas of the test show that some have a greater discriminating capacity than others; those with greater discriminating capacity are: Biology (AUC=0.61, 95% CI from 0.57 to 0.66); Physics/Mathematics (AUC=0.63, 95% CI from 0.59 to 0.68); and Chemistry (AUC=0.65, 95% CI from 0.61 to 0.69) (Table 4).

Once again, the admission test score appears to be predictive of RS2 achievement (AUC= 0.70, 95% CI from 0.64 to 0.76) (Table 4). Also in this case, the same scientific sub-areas have a greater discriminating capacity than the others: Biology (AUC=0.62, 95% CI from 0.56 to 0.68), Physics/Mathematics (AUC=0.63, 95% CI from 0.57 to 0.69) and Chemistry (AUC=0.63, 95% CI from 0.57 to 0.69) (Table 4).

Table 4. Area under ROC curve (AUC) and relative 95% Confidence Intervals for total and per subject area scores

	RS1 *AUC (95%CI)*	**RS2** *AUC (95%CI)*
TOTAL	0.67 (from 0.63 to 0.71)	0.70 (from 0.64 to 0.76)
Biology	0.61 (from 0.57 to 0.66)	0.62 (from 0.56 to 0.68)
Chemistry	0.65 (from 0.61 to 0.69)	0.63 (from 0.57 to 0.69)
General culture	0.55 (from 0.50 to 0.59)	0.57 (from 0.51 to 0.64)
Physics&Mathematics	0.63 (from 0.59 to 0.68)	0.63 (from 0.57 to 0.69)
Logic	0.54 (from 0.50 to 0.60)	0.59 (from 0.53 to 0.65)

Specific analysis of the ROC curve numerical values shows a value between 30 and 35 to be a good cut-off score to use as a selection criterion for student admittance onto the course, in accordance with the institutional one used by the Turin colleges (33.9 in the year 2014-2015). In order to understand the reliability of the test, the real classification used at the University of Turin,

i.e., Regular and TAR students (based on a cut-off score equal to 33.9), was evaluated in relation to Reference Standard RS1 and RS2 (Table 5).

For both RS1 and RS2, Sensitivity is high (Se = 0.86, 95% CI from 0.82 to 0.89 for RS1; Se = 0.91, 95% CI from 0.84 to 0.95 for RS2), whereas Specificity is low (Sp = 0.31, 95% CI from 0.26 to 0.36 for RS1; Sp = 0.25, 95% CI from 0.21 to 0.28 for RS2). Analogously, positive Likelihood ratios LH+ are both higher than 1 (1.23, 95% CI from 1.15 to 1.32 for RS1; 1.22, 95% CI from 1.14 to 1.25 for RS2), whereas negative Likelihood ratios LH- are both less than 1 (0.46, 95% CI from 0.22 to 0.54 for RS1; 0.36, 95% CI from 0.23 to 0.56 for RS2).

The high values of Se indicate that the majority of students who achieve the RSs are from the Regular group, while a low Sp indicates that a number of 'Regular' students are also among those who did not achieve the two RSs (see Table 5). This interpretation is also supported by the good value of LH+, which also indicates that being 'Regular' is a 'protective factor' to achieving both RS1 and RS2.

For our study, however, it is essential to highlight the low value of LH-, which indicates that relatively few TAR students achieved both RSs (in particular RS2). This indicates that a low score in the test also predicts a low probability of achieving both RSs.

Table 5. Predictive capacity of the admission test on the number of CFU acquired

	RS1		**RS2**	
	P1*	N1**	$P2^{+}$	$N2^{++}$
REGULAR	308	223	101	430
TAR	52	100	10	142
TOTAL	360	323	111	572

Se (95% CI)	0.86 (from 0.82 to 0.89)	0.91 (from 0.84 to 0.95)
Sp (95%CI)	0.31 (from 0.26 to 0.36)	0.25 (from 0.21 to 0.28)
LH+ (95% CI)	1.23 (from 1.15 to 1.32)	1.22 (from 1.14 to 1.29)
LH- (95% CI)	0.46 (from 0.22 to 0.54)	0.36 (from 0.23 to 0.56)

**P1: students who acquired half or more of the required credits at the end of the first year*

***N1: students who acquired less than half of the required credits at the end of the first year*

$^{+}$P2: students who acquired all the CFU required for the first year

$^{++}$N2: students who did not acquire all the CFU in the first year

4.DISCUSSION

The admission of a significant number of students to the degree course in medicine at the University of Turin who did not pass the admission test in the year 2014-2015 gave us, for the first time, the opportunity to compare the academic results of two student groups in the same context. A close correlation between academic success and test results was found; in particular, students who got a high score in scientific questions were more likely to achieve the requested standards during the first year. Thus, although improvements could still be made to the admission test, it seems to be a good tool for identifying the students who are more likely to perform well during the first year of the course.

Other studies have analysed the performance of students who passed an admission test versus open access students (see, for example, Reibnegger G. *et al.*[1]), but the comparison could only be made between students of different years. In our case, registration year, lessons, programmes, teachers and classroom characteristics were the same for both Regular and TAR students. For the purposes of our study, it is important to point out that 80% of the TAR students achieved a score of 20 to 31 points in the test (the minimum score for the admission to Turin's School of Medicine was 33.9). This substantiates the TAR group as a good reference.

The Italian test is comparable to other tests used in international contexts, at least in relation to question type and exam duration (Table 1), and for this reason our results may also be of interest outside Italy.

It is important to highlight the similar results found in terms of usefulness and predictability. For example, Sartania N. *et al.* [3] underlined a clear correlation between 'total science score' (our Biology, Chemistry and Physics/Mathematic scores) and 'education performance' (our RS2, although in Sartania N. *et al.* [3] a longer observational period was considered).

Most of the previous studies [1–3] confirmed that admission tests are able to predict the academic results in the first year. Nevertheless, no generalized predictability is assessed in these studies, also because of the variety of the evaluation periods and reference standards. For example, Sartania N. *et al.*[3] evaluated the overall career of the students, while Reibnegger G. *et al.*[1] and Kelly, M. E. *et al.*.[2] and our study considered the results of the first or second years. As far as the reference standards are concerned, Reibnegger G. *et al*[1] investigated the dropout rates (lower in the students passing the admission test), whereas [2] was interested in prediction criteria for clinical and communication skills.

The strengths of this study are the presence of a valid control cohort (TAR) and the possibility of specific analysis per admission test subject area. Although several debates are ongoing in Italy regarding which specific subject areas are most useful for discriminating between potential

medical students, our study shows that the results for questions on biology, chemistry and physics/mathematics in the current admission test present the best predictability.

Several limitations of our study should, however, be taken into consideration. First of all, we only considered students admitted to the course of Medicine in Turin and not a wider Italian cohort. Second, this work constitutes an initial explorative analysis, limited to the first year of the medical course.

It is worth pointing out that the admission test is predictive of academic success, but not necessarily the ability to practice as a Phisician. This aspect has been previously stressed by a number of authors[20–22], but no definitive conclusion has been reached. In general, however, we can say that a single admission test is unable to predict ability to practice as a physician.

Important information could be obtained by following our two cohorts over a longer period of time. Indeed, this study is ongoing in order monitor the two cohorts throughout the complete academic path, re-evaluating their results also with respect to subjects (e.g. clinically oriented courses) different to those considered in the admission test.

Conclusions

In conclusion, university admission test scores are able to predict subsequent academic success in the first year of the degree course in medicine; the test is therefore useful for both students and medical schools. Indeed, it discourages students who do not pass the test from enrolling in the course, driving them towards alternative courses, and saving them both time and money. With this selection procedure in place, universities are able to manage a lower number of more motivated students with higher probabilities of obtaining success. This allows a more efficient use of infrastructural and personnel resources. However, the discriminatory capacity of the admission test could be improved by replicating the analysis presented at the end of the fourth and sixth years, investigating the relationship between admission test results and clinical skills.

Contributorship

All authors contributed toward study design, data analysis and discussion and the drafting of the final manuscript. Specific contributions are as follows:

Giuseppe Migliaretti - design, analysis and discussion and drafting of the final document

Salvatore Bozzaro – discussion and drafting of the final document

Roberta Siliquini - discussion and drafting of the final document

Giuseppe Costa – design, analysis and discussion and drafting of the final document

Franco Cavallo - design, analysis and discussion and drafting of the final document

Acknowledgements

We wish to thank for their valued support:

Prof Anna Bossi – Dept of Clinical Sciences and Community, University of Milan

Prof Anita Tabacco – DISMA Dept of Mathematical Sciences "G. L. Lagrange", Turin Polytechnic

Prof Caterina Guiot – Dept of Neurosciences, University of Turin

Dr Barbara Martin – Dept of Chemistry, University of Turin

Dr. Paola Ropolo – Student services of the School of Medicine, University of Turin

Competing interests: no competing interests are present

Funding: the work has not received any kind of economic support

References

1. Reibnegger, G. *et al.* Dropout rates in medical students at one school before and after the installation of admission tests in Austria. *Acad. Med. J. Assoc. Am. Med. Coll.* **86,** 1040–1048 (2011).
2. Kelly, M. E. *et al.* To what extent does the Health Professions Admission Test-Ireland predict performance in early undergraduate tests of communication and clinical skills? An observational cohort study. *BMC Med. Educ.* **13,** 68 (2013).
3. Sartania, N., McClure, J. D., Sweeting, H. & Browitt, A. Predictive power of UKCAT and other pre-admission measures for performance in a medical school in Glasgow: a cohort study. *BMC Med. Educ.* **14,** 116 (2014).
4. Prideaux, D. *et al.* Assessment for selection for the health care professions and specialty training: consensus statement and recommendations from the Ottawa 2010 Conference. *Med. Teach.* **33,** 215–223 (2011).
5. Segouin, C. *et al.* Country report: medical education in France. *Med. Educ.* **41,** 295–301 (2007).
6. Coebergh, J. Dutch medical schools abandon selection for lottery system for places | Student BMJ. **11,** 138
7. Decreto Ministeriale del 29 novembre 2013 n. 986 - Atti Ministeriali MIUR. Available at: http://attiministeriali.miur.it/anno-2013/novembre/dm-29112013.aspx. (Accessed: 24th July 2017)
8. Al-Alwan, I. A. Association between scores in high school, aptitude and achievement exams and early performance in health science college. *Saudi J. Kidney Dis. Transplant. Off. Publ. Saudi Cent. Organ Transplant. Saudi Arab.* **20,** 448–453 (2009).
9. Yates, J. & James, D. The UK Clinical Aptitude Test and clinical course performance at Nottingham: a prospective cohort study. *BMC Med. Educ.* **13,** 32 (2013).
10. Parry, J. *et al.* Admissions processes for five year medical courses at English schools: review. *BMJ* **332,** 1005–1009 (2006).
11. Delavar, M. A. *et al.* Using the objective structured clinical examinations in undergraduate midwifery students. *J. Med. Life* **6,** 76–79 (2013).
12. Berg, K. *et al.* The relationship between performance on a medical school's clinical skills assessment and USMLE Step 2 CS. *Acad. Med. J. Assoc. Am. Med. Coll.* **83,** S37-40 (2008).

13. Lancia, L., Petrucci, C., Giorgi, F., Dante, A. & Cifone, M. G. Academic success or failure in nursing students: results of a retrospective observational study. *Nurse Educ. Today* **33,** 1501–1505 (2013).

14. Mannella, R. The Italian veterinary medicine admission test: analysis of student intake in the years 2007, 2008, and 2009, and of the test's relationship with students' academic careers. *J. Vet. Med. Educ.* **38,** 184–193 (2011).

15. Altman, D. *Practical statistics for medical research*. (Chapman & Hall).

16. Hanley, J. A. & McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143,** 29–36 (1982).

17. Zweig, M. H. & Campbell, G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.* **39,** 561–577 (1993).

18. Harper, R. & Reeves, B. Reporting of precision of estimates for diagnostic accuracy: a review. *BMJ* **318,** 1322–1323 (1999).

19. Nam, J. □M. Confidence Limits for the Ratio of Two Binomial Proportions Based on Likelihood Scores: Non□Iterative Method. *Biom. J.* **37,** 375–379 (1995).

20. Louridas, M., Szasz, P., de Montbrun, S., Harris, K. A. & Grantcharov, T. P. Can We Predict Technical Aptitude?: A Systematic Review. *Ann. Surg.* **263,** 673–691 (2016).

21. Kreiter, C. D. & Axelson, R. D. A perspective on medical school admission research and practice over the last 25 years. *Teach. Learn. Med.* **25 Suppl 1,** S50-56 (2013).

22. Patterson, F. *et al.* How effective are selection methods in medical education? A systematic review. *Med. Educ.* **50,** 36–60 (2016).

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract **OK (pag.1-2)**	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported **OK (pag.4-5 and pag.6 rows 1-6)**
Objectives	3	State specific objectives, including any prespecified hypotheses **OK (pag.6 rows 8-11)**
Methods		
Study design	4	Present key elements of study design early in the paper **OK (pag.6 rows 27-32 and pag.7 rows 1-24 and Ref. 15-19)**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **OK (pag.6 rows 19-31)**
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls **OK (pag.5 rows 26-31 and pag.6 rows 19-23)** *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable **OK (pag.6 rows 19-31)**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **OK (pag.6 rows 13-31 and pag.7 rows 1-24)**
Bias	9	Describe any efforts to address potential sources of bias **OK (pag.11 rows 16-27)**
Study size	10	Explain how the study size was arrived at **OK (pag.5 rows 26-31 and pag.6 rows 19-25)**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why **OK (pag.6 rows 28-31)**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding **OK (pag.7 rows 1-24)**
		(*b*) Describe any methods used to examine subgroups and interactions **not appl**
		(*c*) Explain how missing data were addressed **NO Missing data**
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Results		
Participants **OK (pag.5 rows 26-31 and pag.6 rows 19-23)**	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders **OK (Table 2 and 3)**
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure **OK (pag.6 rows 19-23 and Table 2 and Table 3)**
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results **OK (pag.7 rows 26-31 and pag.8-9)**	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses **OK**	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives **OK (pag.10 rows 19-25)**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias **OK (pag.11 rows16-28)**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence **OK (pag.10 rows 29-33)**
Generalisability	21	Discuss the generalisability (external validity) of the study results **OK (pag.10 rows 34-36 and pag.11 rows1-18)**
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based **Not appl**

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

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