Selected medical students achieve better than lottery-admitted students during clerkships

Louise C Urlings-Strop, Axel P N Themmen, Theo Stijnen & Ted A W Splinter

In: Urlings-Strop LC, Themmen AP, Stijnen T, Splinter TA. Selected medical students achieve better than lottery-admitted students during clerkships. Med Educ. 2011;45:1032-1040.

ABSTRACT

Objectives – A recent controlled study by our group showed that the dropout rate in the first 2 years of study of medical students selected for entry by the assessment of a combination of non-cognitive and cognitive abilities was 2.6 times lower than that of a control group of students admitted by lottery. The aim of the present study was to compare the performance of these two groups in the clinical phase.

Methods – A prospective cohort study was performed to compare the performance of 389 medical students admitted by selection with that of 938 students admitted by weighted lottery between 2001 and 2004. Follow-up of these cohorts lasted 5.5–8.5 years. The main outcome measures were the mean grade obtained on the first five discipline-specific clerkships by all cohorts and the mean grade achieved on all 10 clerkships by the cohorts of 2001 and 2002.

Results – Selected students obtained a significantly higher mean grade during their first five clerkships than lottery-admitted students (mean ± standard error [SE] 7.95 ± 0.03, 95% confidence interval [CI] 7.90–8.00 versus mean ± SE 7.84 ± 0.02, 95% CI 7.81–7.87; $P < 0.001$). This difference reflected the fact that selected students achieved a grade of ≥ 8.0 1.5 times more often than lottery-admitted students. An analysis of all mean grades awarded on 10 clerkships revealed the same results. Moreover, the longer follow-up period over the clerkships showed that the relative risk for dropout was twice as low in the selected student group as in the lottery-admitted student group.

Conclusions – The selected group received significantly higher mean grades on their first five clerkships, which could not be attributed to factors other than the selection procedure. Although the risk for dropout before the clinical phase increased somewhat in both groups, the actual dropout rate proved to be twice as low in the selected group.

INTRODUCTION

The number of applicants to medical school exceeds the number of places available. Therefore, student selection is an internationally widespread practice. Various methods are used to select the best performing students. The most common methods are the undergraduate grade point average (uGPA) and the Medical College Admission Test (MCAT) (Medical School Admission Requirements (MSAR) 2008–2009: United States and Canada, 2007; Parry, Mathers, et al., 2006). The uGPA has the greatest bearing on student achievement in medical school (Baars, Wimmers, et al., 2009; Cohen-Schotanus, Muijtjens, et al., 2006). Most studies suggest that it explains 15–25% of variance in achievement during the pre-clinical phase (Baars, Wimmers, et al., 2009; Salvatori, 2001).

However, results have shown that it is much more difficult to predict future clinical performance during clerkships (Basco Jr., Gilbert, et al., 2000; Hamdy, Prasad, et al., 2006). Given an explained variance of < 10%, the relationship between uGPA and achievement during clerkships is much weaker (Baars, Wimmers, et al., 2009; Peat, Woodburry, et al., 1982; Veloski, Callahan, et al., 2000). Most probably, the decrease in the strength of the relationship is caused by the substantial change in the learning environment – including in methods of assessment – at the transition from the pre-clinical to the clinical phase (Prince, Van De Wiel, et al., 2000). Similarly, the MCAT has a moderate to low predictive value for clerkships, with an explained variance in achievement of approximately 15% (Donnon, Paolucci, et al., 2007).

Both the uGPA and MCAT represent cognitive domains. Of non-cognitive measures, the most frequently used is the interview, sometimes accompanied by letters of reference or psychological tests, although predictive validity correlations rarely rise above 0.10 (Albenese, Snow, et al., 2003; Salvatori, 2001). To enhance the predictive validity of the interview, the multiple mini-interview (MMI) was developed; this is an admission procedure resembling an objective structured clinical examination with multiple short-interview stations. Although the results of only small groups of students have been presented, the MMI has shown promising results in predicting clerkship and licensing examination results (Reiter, Eva, et al., 2007). It is therefore unsurprising to identify a need for other criteria, especially those that are non-cognitive, with which to assess personal qualities for the purposes of selecting medical students (Collin, Violato, et al., 2009; Eva, Reiter, et al., 2004; Ferguson, James, et al., 2002).

Since 2000, medical schools in the Netherlands have been allowed to select up to 50% of their students based on characteristics other than pre-university GPA (pu-GPA). ten Cate & Hendrix (2001) and ten Cate, Hendrix, et al. (2002) have described the selection procedures used in two medical schools. In one of these, selection was based on cognitive abilities; in the other, selection relied on assessment of non-cognitive abilities such as motivation and views on the medical profession. An analysis carried out after 1 year of follow-up found that selected students obtained a slightly higher mean grade on written examinations. Both medical schools selected 24 students, whose results were compared with those of 341 lottery-admitted students. Hulsman, van der Ende, et al. (2007) used a three-step procedure. They selected 56 students and compared their outcomes with those of 446 lottery-admitted controls in two cohorts. The selection procedure referred to in this study involved the writing of an essay, cognitive tests and an examination of social skills. After 1 year of follow-up, no difference in academic achievement was found between selected students and controls.

From 2001, we performed a controlled experiment to find out whether students selected for medical school using a combination of cognitive and non-cognitive assessments would achieve more than those selected by lottery. The first four cohorts consisted of a total of 389 selected students and 938 lottery-admitted controls. In short, the main outcome of this prior study was the relative risk for dropout in the first 2 years, which we found to be 2.6 times lower for selected students than it was for students admitted by lottery (Urlings-Strop, Stijnen, et al., 2009). The four cohorts of selected and lottery-admitted students who entered medical school between 2001 and 2004 were followed to track their achievements in the clinical phase. The present paper reports the comparison of their clinical achievements.

The existence of the lottery-admitted group next to the selected group provided a unique opportunity to compare the results of randomly admitted versus selected students. This is the first study to compare the clinical achievements of selected students and controls.

METHODS

Selection procedure

Since 2000, there have been three ways of gaining admittance to medical school in the Netherlands: a local selection procedure (S-group); the national lottery system (L-group), and an unrestricted direct access (D-group). Applicants taking part in the

lottery can also choose to apply for a local selection procedure, which precedes the lottery. The local selection procedure at Erasmus MC medical school and the weighted lottery system in the Netherlands have already been described (Urlings-Strop, Stijnen, et al., 2009). Direct access is given to students with a pu-GPA of ≥ 8.0.

Clinical curriculum

The medical curriculum consists of a 4-year (168-week) pre-clinical phase followed by a 2-year (84-week) clinical phase. The clinical phase consists of two parts which involve a fixed sequence of clinical rotations. A period of 15 weeks of general clinical training precedes 69 weeks of discipline-specific clerkships. The general clinical training includes 4 weeks of introduction followed by three short clerkships in internal medicine, surgery and paediatrics of 4, 4 and 3 weeks, respectively. Student achievement in each of these three clerkships is judged as 'insufficient', 'sufficient' or 'good'. The aim of the general clinical training is to teach the student the following competencies: history taking; physical examination; differential diagnosis; additional diagnostic procedures, and therapy planning (Splinter & Verwoerd, 2000). Because of the non-discriminatory power of this grading, these results were not included in further analysis.

Discipline-specific clerkships include 11 different rotations in internal medicine, surgery, paediatrics, psychiatry, neurology, gynaecology, dermatology, otorhinolaryngology, ophthalmology, family medicine and public health, respectively, followed by 15 weeks of elective clerkships. However, performance in both the public health and elective clerkships is not graded. The durations of the 10 clerkships on which performance is graded are 8, 8, 3, 5, 5, 5, 3, 3, 3 and 4 weeks, respectively. Some of these are preceded by a week of introductory courses. At the end of each clerkship, the student's performance is assessed using a combination of patient-related assessment and oral examination and rewarded with a grade between 5 (unsatisfactory) and 10 (excellent). In addition, overall clinical performance is assessed by a supervisor during each clerkship.

Participants

Students are eligible to start the clinical phase when they have completed all obligatory modules and all examinations. On the date selected to qualify for entry to this study (1 April 2010) 78.3% of students in the 2001 and 2002 cohorts had passed all 10 clerkships and 65.8% of those in the 2001–2004 cohorts had completed the first five discipline-specific clerkships. Students who were not eligible to start the clinical phase and who had obtained no grades in the pre-clinical phase for more than a year were considered as late dropouts; students who had failed to complete the

first year of study within 2 years were considered as early dropouts. The mean grade for the 10 discipline-specific clerkships in the cohorts of 2001 and 2002 and that for the first five clerkships in all four cohorts was used as the criterion representing clinical achievement.

Statistical analysis

For all students, pre-admission data on age, gender and pu-GPA were collected, as well as start date and the number of clerkships completed by 1 April 2010. Within each cohort, pu-GPAs were translated into Z-scores in order to diminish cohort differences. Data on gender and the percentage of students who had begun clerkships, as well as on overall dropout rate and grades obtained on the clerkships, were analysed using the chi-squared test or Mantel–Haenszel test with stratification by year of entrance and weighted lottery category if appropriate. Analysis of covariance (ANCOVA) was used for comparisons of grades obtained. Again, year of entry and weighted lottery category were used as covariates.

All data were derived from the university student administration system, recorded in EXCEL 2003 workbooks and analysed using PASW for Windows Version 17.0 (SPSS, Inc., Chicago, IL, USA).

Ethical considerations

In the Netherlands, ethical approval for studies concerning medical education is not required. However, in order to adhere to the tenets of the Declaration of Helsinki, we took some precautions. Members of the medical faculty, who were not involved in this research, were authorised by the examination committee to extract the students' grades from the university administration system and deliver them in an anonymised form to the investigators.

RESULTS

As 65.8% of the 2001–2004 cohorts had completed at least five clerkships and 78.3% of the 2001 and 2002 cohorts had completed all 10 clerkships, the mean grades for five and 10 clerkships were used to compare the clinical achievement of selected and lottery-admitted students. An almost ideal correlation of 0.89 was found between the mean grade achieved on the first five clerkships and the mean grade obtained on all 10 clerkships by the 443 students in the cohorts of 2001 and 2002. This indicated that former achievement could be considered representative of overall clinical achievement. For the purposes of clarity, data for both the four

cohorts of 2001–2004 and the two cohorts of 2001 and 2002 are shown in Tables 1–3, but only data for the four cohorts of 2001–2004 are discussed.

Table 1 summarises quantitative data from the start of medical education up to the completion of either five or 10 clerkships. In the four cohorts of 2001– 2004, 389 students were selected (S-group) and 938 students were admitted by lottery (L-group).

Table 1 – *Quantitative characteristics of the cohorts of 2001-2004 and the cohorts of 2001 and 2002 after 5.5 - 8.5 years of follow-up, respectively*

	Cohorts of 2001-2004				Cohorts of 2001-2002			
	S-group		L-group		S-group		L-group	
	n	(%)	*n*	(%)	*n*	(%)	*n*	(%)
Started medical school	389		938		103		463	
Early drop-out in first 2 years	24	(6.2)	140	(14.9)	4	(3.9)	72	(15.6)
Qualified to complete pre-clinical phase	365		798		99		391	
Late drop-out	12	(3.1)	20	(2.1)	0	(0.0)	10	(2.2)
Not yet completed pre-clinical phase	15	(3.9)	23	(2.5)	1	(1.0)	0	(0.0)
Eligible to start clerkships	338	(86.9)	755	(80.5)	98	(95.1)	381	(82.3)
Completed pre-clinical but stopped	16	(4.1)	27	(2.9)	2	(1.9)	3	(0.6)
Completed < 5* / 10† clerkships	80	(20.6)	97	(10.3)	9	(8.7)	22	(4.8)
Completed ≥ 5* / 10† clerkships	242	(62.2)	631	(67.2)	87	(84.5)	356	(76.9)

* For the cohorts of 2001-2004
† For the cohorts of 2001 and 2002
S-group = selected students; L-group = lottery admitted students

Of the 365 selected students and 798 lottery-admitted students who completed the pre-clinical phase, 12 and 20 students, respectively, dropped out in subsequent years, although they completed the first year, which led to overall dropout rates of 9.3% in the S-group and 17.1% in the L-group. After controlling for cohort and weighted lottery category, this differ- ence remains significant ($\chi^2_{(1)} = 10.11$, $p < 0.001$). The risk for both early and late dropout during the entire pre-clinical phase remained twice as large in the L-group as it was in the S-group.

Despite a follow-up of ≥ 5.5 years, 15 students in the S-group (3.9%) and 23 in the L-group (2.5%) had not yet completed the pre-clinical phase. The difference between the groups is not significant. Having completed the pre-clinical phase, 338 students in the S-group (86.9%) and 755 students in the L-group (80.5%) were eligible to start the clerkships. Sixteen (4.1%) students in the former group and 27 (2.9%) in the latter did not actually start this phase. Of these 16 S-group students,

13 were involved in a PhD programme; of the 27 L-group students, 17 were enrolled in PhD programmes, one was studying law and one was studying public health. No such activities could be confirmed for the 11 remaining students.

Eighty (20.6%) students in the S-group and 97 (10.3%) in the L-group in the four cohorts of 2001– 2004 had started the clerkships but had not yet completed five of them. Thus, data on the mean grade achieved on five clerkships were available for 242 (62.2%) students in the S-group and 631 (67.2%) in the L-group. Using the Mantel–Haenszel stratification test (strata = year of entrance), the difference between the groups in the percentage of students at this stage proved non-significant ($\chi^2_{(1)} = 0.23$, $p > 0.05$). The percentage of students who were eligible to start clerkships but had not completed at least five discipline-specific clerkships was 24.7% in the S-group and 13.2% in the L-group ($\chi^2_{(1)} = 2.97$, $p < 0.001$).

Pre-admission variables in the S- and L-groups in successive stages of study are shown in Table 2. There is a significant difference between the groups in the percentage of women at all stages, except at the start of medical school, when comparing gender over all four cohorts. Between the start of medical school and the completion of five clerkships, the percentage of women increased by 6.5% in the S-group and 4.2% in the L-group. These data indicate the positive selection of women during the pre-clinical phase. The variable 'mean age at start' remained the same during this period. However, the S-group was significantly older than the L-group, although the difference was usually < 6 months.

No significant difference in pu-GPA was observed between the S- and L-groups at any stage. Pre- university GPA increased slightly from the start through the five clerkships, indicating a positive selection during the pre-clinical phase. Because of the significant difference in gender and age, these variables were added as a covariate to the ANCOVA when the mean grades of the clerkships were assessed.

During the clinical phase, a total of 4365 grades were given for the first five clerkships. For administrative reasons, 40 grades (0.9%) were missing. An ANCOVA was used to compare the mean grades obtained by both groups on these first five clerkships, taking into account the cohort in which the students had started their study and their weighted lottery category. The first variable had no significant bearing on the mean grade, but the latter was significantly related ($F_{(1,827)} = 27.17$, $p < 0.001$). Mean ± standard error (SE) grades were 7.95 ± 0.03 (95% confidence interval [CI] 7.90–8.00) in the S-group and 7.84 ± 0.02 (95% CI 7.81–7.87) in the L-group ($F_{(1,822)} = 12.30$, $p < 0.001$). The effect size for this mean difference in GPA was small (0.015).

Table 2 – *Pre-admission variables at the start of medical school and after completion of at least five clerkships for the cohorts 2001-2004 and of 10 clerkships for the cohorts of 2001 and 2002*

		Cohorts of 2001-2004				Cohorts of 2001 and 2002			
		At start of medical school		Finished ≥ 5 clerkships		At start of medical school		Finished 10 clerkships	
Gender, % female									
	S-group	65.0		71.5†		73.8‡		78.0§	
	L-group	60.5		64.7		59.8		64.1	
Mean Age, years (SD)									
	S-group	19.69¶	(0.09)	19.74**	(0.12)	20.01††	(0.18)	20.02‡‡	(0.18)
	L-group	19.34	(0.04)	19.32	(0.07)	19.52	(0.08)	19.37	(0.06)
Mean pu-GPA* (SD)									
	S-group	-0.17	(0.01)	-0.04	(0.02)	-0.17	(0.03)	-0.09	(0.03)
	L-group	-0.14	(0.01)	-0.04	(0.01)	-0.15	(0.01)	-0.07	(0.02)

* Z-score
† Differs significantly from corresponding L-group ($\chi^2_{(1)} = 4.01$, $p < 0.05$)
‡ Differs significantly from corresponding L-group ($\chi^2_{(1)} = 7.10$, $p < 0.01$)
§ Differs significantly from corresponding L-group ($\chi^2_{(1)} = 8.24$, $p < 0.01$)
¶ Differs significantly from corresponding L-group ($F_{(1,1259)} = 9.96$, $p < 0.01$)
** Differs significantly from corresponding L-group ($F_{(1,824)} = 9.56$, $p < 0.01$)
†† Differs significantly from corresponding L-group ($F_{(1,532)} = 6.28$, $p < 0.05$)
‡‡ Differs significantly from corresponding L-group ($F_{(1,392)} = 10.32$, $p < 0.001$)
S-group = selected students; L-group = lottery admitted students; SD = standard deviation; pu-GPA = pre-university grade point average

The addition of gender and age as covariates to the ANCOVA when mean grades were assessed showed no difference in means.

To explain the difference in mean grades between the S- and L-groups, the grades were divided into two categories, comprising: (i) grades of > 5.5 to< 8.0, and (ii) grades of ≥ 8.0. The first category represents grades of 6.0 (below average, but just good enough to continue) and 7.0 (average); the second category represents grades of 8.0 (good) and 9.0 (very good). In 7.5% of cases, grades were rounded up to a whole number. Table 3 shows the distribution of grades. Overall, over 70% of the grades were ≥ 8.0. However, S-group students obtained grades of ≥ 8.0 significantly more often than L-group students ($\chi^2_{(1)} = 30.17$, $p < 0.001$). The probability of achieving a grade ≥ 8.0 was 1.5 times higher for selected students than it was for lottery- admitted students.

Table 3 – *Number of students achieving grades of < 8.0 or ≥ 8.0*

	Cohorts of 2001-2004		Cohort of 2001 and 2002	
	S-group*	L-group	S-group†	L-group
	n (%)	n (%)	n (%)	n (%)
5.5 < grade < 8	256 (21.5)	933 (29.8)	174 (20.0)	1043 (29.3)
Grade ≥ 8	936 (78.5)	2200 (70.2)	695 (80.0)	2516 (70.7)

* The S-group obtained a grade of ≥ 8 more often ($\chi^2_{(1)} = 30.17$, $p < 0.001$)
† The S-group obtained a grade ≥ 8 more often ($\chi^2_{(1)} = 30.20$, $p < 0.001$)
In the cohorts of 2001-2004, the mean grade was 7.95 (SE = 0.03, 95% CI = 7.90-8.00) in the S-group and 7.84 (SE = 0.02, 95% CI = 7.81-7.87) for the L-group ($F_{(1,822)} = 12.30$, $p < 0.001$).
In the cohorts of 2001 and 2002, the estimated mean grade is 8.00 (SE = 0.04, 95% CI = 7.93-8.07) in the S-group and 7.84 (SE = 0.02, 95% CI = 7.81-7.87) for the L-group ($F_{(1,414)} = 15.25$, $p < 0.001$).
S-group = selected students; L-group = lottery admitted students; SE = standard error; CI = confidence interval

DISCUSSION

In this study, we report the mean grades of selected (S-group) and lottery-admitted (L-group) students in the cohorts of 2001–2004 on the first five discipline-specific clerkships and in the 2001 and 2002 cohorts on all 10 discipline-specific clerkships. The S-group obtained a higher mean grade than the L-group on the first five as well as all 10 clerkships. The difference is small, yet significant, and reflects the fact that S-group students obtained mean grades of ≥ 8.0 1.5 times more often than L-group students. The mean grade on the first five clerkships of 443 recently graduated students (2001 and 2002 cohorts) showed a high correlation ($r = 0.89$) with the overall mean grade. These data suggest that our findings may be extrapolated to the overall clinical achievement of all students in the four cohorts of 2001–2004. This is an important finding because the literature reports correlations between selection methods such as the MCAT or pu-GPA and subsequent clinical performance as low and often not significant (Salvatori, 2001; White, Dey, et al., 2009).

In order to find out whether the observed difference in clinical achievement was related to selection before admission to medical school, pre-admission variables and selective steps after admission (but before the start of clerkships) were compared between the S- and L-groups. The early dropout rate in the S-group was significantly lower than in the L-group. In addition, although the overall risk for dropout declined, it remained twice as low in the S-group as in the L-group, representing a highly significant difference between both groups. The percentage of students who were still in the process of completing the pre-clinical 4-year phase of education after ≥ 5.5 years was small and did not differ significantly between the two groups. All students

are expected to eventually complete the pre-clinical phase, as did all students in the cohorts of 2001 and 2002.

The percentage of students eligible to start the clerkships was higher in the S-group than in the L-group, mainly as a result of the difference in pre-clinical dropout rates. The difference was not significant. Similarly, the percentage of students who had completed five of 10 clerkships did not differ significantly between the two groups. However, the percentage of students who were eligible to begin clerkships but either stopped or started later and were therefore able to complete fewer than five clerkships before the qualifying date of 1 April 2010 was significantly higher in the S-group. University personnel records identified only those students who had completed the pre-clinical phase and started on a PhD programme or a different course of study at our university. This accounted for only 13 (3.3%) students in the S-group and 19 (2.0%) students in the L-group. Thirty of these had started on a PhD programme and two had embarked on an additional course of study; these numbers were roughly equally divided between the two groups. It was common for students to choose to postpone the start of the clerkships and related commitments to patient care in favour of other activities, such as travel abroad. However, we do not know whether the frequency of such non-study-related activities was higher among selected students than among lottery-admitted students. Nevertheless, it is highly probable that students who postponed the start of the clerkships had completed the pre-clinical phase in an optimal and nominal way.

Comparison of the pre-admission variables of gender, mean age and pu-GPA showed an increase in the percentage of women between the start of medical school and the completion of five or 10 clerkships. The increase was almost identical in both the S- and L-groups and indicated the positive selection of women during the pre-clinical phase. Mean age at the start of medical school remained the same during the period of study and indicated the absence of a relationship between this pre-admission variable and student achievement in the pre-clinical phase. Finally, pu-GPA in the S- and L-groups increased slightly during the period between the start and completion of clerkships, indicating the positive selection of students with a higher pu-GPA. More importantly, no significant difference in pu-GPA was observed between the S- and L-groups at any stage between the start of medical school and graduation after the completion of clerkships. Therefore, the observed difference in clinical achievement between the S- and L-groups appears to be related to the selection of students before admission.

The difference in mean grades on clerkships between the S- and L-groups is small because the distribution of grades was limited to four possibilities: 6.0, 7.0, 8.0 or 9.0. More than 70% of students obtained grades of 8.0 or 9.0. The difference in mean grades was caused by the fact that the probability of achieving a grade of ≥ 8.0 was 1.5 times greater for selected students than for lottery-admitted students. This is a significant and relevant difference.

Grades for clinical achievement in the four cohorts of 2001–2004 were based on subjective global performance ratings (GPRs) during the clerkships and a patient-related plus an oral examination at the end. The GPR represents a global rating awarded by a supervisor and covers the student's performance on a number of clinically relevant competencies over a certain period of time (Daelmans, van der Hem-Stokroos, et al., 2005). Recently, Wimmers, Kanter, et al. (2008) tried to establish which competencies were important for clerkship grading by administering a survey to clinical teachers at 17 teaching hospitals at which our students undertake their clerkships. The survey consisted of items on 21 different student characteristics, which clinical teachers were asked to rank in order of their importance to clerkship grading. Using structural equation modelling, a four-factor structure was found to define a competence profile for clerk- ship grading. The factors were 'cognitive abilities', 'patient workup', 'interpersonal skills' and 'professional qualities'. The first two factors were considered significantly more important by grading teachers than the other two, and indicated that grading was mainly determined by knowledge, the quality of the patient file and problem-solving abilities (Wimmers, Kanter, et al., 2008). Unfortunately, a more precise description of how much of the grading reflected pure cognitive or non-cognitive accomplishments cannot be given. This is an assessment method used not only in the Netherlands, but also worldwide (Kasselbaum & Eaglen, 1999) and, despite efforts to implement more modern forms of assessment, clinical examinations remain of questionable reliability and validity (van der Vleuten, 2000). It should be kept in mind that both groups sat the same examinations and supervisors had no knowledge of whether a student had been selected or lottery-admitted.

These data may be limited by the fact that not all eligible students in the four cohorts had completed the clinical phase. However, as the difference in mean clinical grade between the S- and L-groups was also found between the students in the cohorts of 2001 and 2002, most of whom had completed 10 clerkships, and as the mean grade on the first five clerkships was highly correlated (r = 0.89) to the mean grade on all 10 clerkships, the addition of more results will probably not change the outcome in any significant manner.

Another possible limitation is the subjective assessment method. Subjectivity is the main reason why the grades are abnormally distributed so that > 70% of students obtained grades in the highest categories of 8.0 and 9.0. Such grades are much higher than the pu-GPA (6.9) and the mean grade in the pre-clinical phase of training (6.1) achieved by the same cohort of students. The abnormal distribution limits distinction among the clinical competencies of the students and results in a small difference between the mean grades of the S- and L-groups. Both groups were subject to the same supervisors, global performance ratings and examinations. Supervisors and examiners were completely unaware of whether a student had been selected or lottery-admitted.

In conclusion, follow-up periods of 5.5–8.5 years of selected and lottery-admitted students in the cohorts of 2001–2004 revealed that selected students obtained a significantly higher mean grade on clerk- ships than lottery-admitted students. This difference in mean grade probably indicates differences in knowledge, quality of patient files and problem-solving abilities. The observed difference appears to relate to selection before admission. Reports on factors that have some predictive value for clinical achievement are rare (Donnon, Paolucci, et al., 2007; Meredith, Dunlap, et al., 1982; Reiter, Eva, et al., 2007). Our findings in the controlled setting of the national lottery system represent a strong stimulus for further research on other characteristics, which may select students with specific clinical competencies. However, such an effort is only worthwhile if such competencies are valued, taught and measured in medical school. For such research, we must first investigate the existence of any cause–effect relationship among our selection methods.

REFERENCES

Albenese MA, Snow MH, Skochelak SE, K.N. H and Farrell PM. Assessing personal qualities in medical school admissions. *Acad Med.* 2003;**78**(3):313-321.

Baars GJA, Wimmers PF, Stijnen T, Schmidt HG and Splinter TAW. A model to predict student failure in the first-year medical curriculum [PhD-thesis]. Rotterdam: Erasmus University. 2009.

Basco Jr. WT, Gilbert GE, Chessman AW and Blue AV. The ability of a medical school admission process to predict clinical performance and patients' satisfaction. *Acad Med.* 2000;**75**(7):743-747.

Cohen-Schotanus J, Muijtjens AM, Reinders JJ, Agsteribbe J, van Rossum HJ and van der Vleuten CP. The predictive validity of grade point average scores in a partial lottery medical school admission system. *Med Educ.* 2006;**40**(10):1012-1019.

Collin VT, Violato C and Hecker K. Aptitude, achievement and competence in medicine: a latent variable path model. *Adv Health Sci Educ.* 2009;**14**(3):355-366.

Daelmans HE, van der Hem-Stokroos HH, Hoogenboom RJ, Scherpbier AJ, Stehouwer CD and van der Vleuten CP. Global clinical performance rating, reliability and validity in an undergraduate clerkship. *Neth J Med.* 2005;**63**(7):279-284.

Donnon T, Paolucci EO and Violato C. The predictive validity of the MCAT for medical school performance and medical board licensing examinations: a meta-analysis of the published research. *Acad Med.* 2007;**82**(1):100-106.

Eva KW, Reiter HI, Rosenfeld J and Norman GR. The ability of the multiple mini-interview to predict preclerkship performance in medical school. *Acad Med.* 2004;**79**(10, suppl):40-42.

Ferguson E, James D and Madeley L. Factors associated with success in medical school: systematic review of the literature. *BMJ.* 2002;**324**(7343):952-957.

Hamdy H, Prasad K, Anderson MB, Scherpbier A, Williams R, Zwierstra R and Cuddihy H. BEME systematic review: Predictive values of measurements obtained in medical schools and future performance in medical practice. *Med Teach.* 2006;**28**(2):103-116.

Hulsman RL, van der Ende JSJ, Oort FJ, Michiels RPJ, Casteelen G and Griffioen FMM. Effectiveness of selection in medical school admissions: evaluation of the outcomes among freshmen. *Med Educ.* 2007;**41**:369-377.

Kasselbaum DG and Eaglen RH. Shortcomings in the evaluation of students' clinical skills and behaviors in medical school. *Acad Med.* 1999;**74**(7):842-849.

Medical School Admission Requirements (MSAR) 2008–2009: United States and Canada. 57 ed. Washington: Association of American Medical Colleges. 2007.

Meredith KE, Dunlap MR and Baker HH. Subjective and objective admissions factors as predictors of clinical clerkship performance. *J Med Educ.* 1982;**57**(10):743-751.

Parry J, Mathers J, Stevens A, Parsons A, R L, Spurgeon P and Thomas H. Admissions processes for five year medical courses at English schools: review. *BMJ.* 2006;**332**:1005-1009.

Peat M, Woodburry MG and Donner A. Admission average as predictor of undergraduate academic and clinical performance. *Physiother Can.* 1982;**34**:211-214.

Prince KJ, Van De Wiel M, Scherpbier AJ, Can Der Vleuten CP and Boshuizen HP. A Qualitative Analysis of the Transition from Theory to Practice in Undergraduate Training in a PBL-Medical School. *Adv Health Sci Educ.* 2000;**5**(2):105-116.

Reiter HI, Eva KW, Rosenfeld J and Norman GR. Multiple mini-interviews predict clerkship and licensing examination performance. *Med Educ.* 2007;**41**(4):378-384.

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

Splinter TAW and Verwoerd CDA. De Erasmus-arts 2007, voorstel voor een nieuw curriculum (The Erasmus-doctor 2007, proposal for a new curriculum). Rotterdam: Erasmus Universiteit Rotterdam. 2000.

ten Cate TJ and Hendrix HL. De eerste ervaringen met selectie voor de artsopleiding. *Ned Tijdschr Geneeskd.* 2001;**145**(28):1364-1368.

ten Cate TJ, Hendrix HL, de Fockert Koefoed KJJ and Rietveld WJ. Studieresultaten van toegelatenen binnen en buiten de loting. *Tijdschr Med Ond.* 2002;**21**(6):253-258.

Urlings-Strop LC, Stijnen T, Themmen AP and Splinter TA. Selection of medical students: a controlled experiment. *Med Educ.* 2009;**43**(2):175-183.

van der Vleuten C. Validity of final examinations in undergraduate medical training. *BMJ.* 2000;**321**(7270):1217-1219.

Veloski JJ, Callahan CA, Xu G, Hojat M and Nash DB. Prediction of students' performance on licensing examinations using age, race, sex, undergraduate GPAs and MCAT scores. *Acad Med.* 2000;**75**(10, suppl):28-30.

White CB, Dey EL and Fantone JC. Analysis of factors that predict clinical performance in medical school. *Adv Health Sci Educ.* 2009;**14**(4):455-464.

Wimmers PF, Kanter SL, Splinter TAW and Schmidt HG. Is clinical competence perceived differently for student daily performance on the wards versus clerkship grading? *Adv Health Sci Educ.* 2008;**13**(5):693-707.