

Lessons learned from 15 years of non-grades-based selection for medical school

Karen M Stegers-Jager 

CONTEXT Thirty years ago, it was suggested in the Edinburgh Declaration that medical school applicants should be selected not only on academic, but also on non-academic, attributes. The main rationale behind extending medical school selection procedures with the evaluation of (non-academic) personal qualities is that this will lead to the selection of students who will perform better as a doctor than those who are selected on the basis of academic measures only. A second rationale is the expectation that this will lead to a representative health workforce as a result of reduced adverse impact. The aims of this paper are (i) to describe what can be learned about the use of selection criteria other than grades from over 15 years of Dutch experience and (ii) to summarise current knowledge on the issue of adverse impact in relation to non-grades-based selection.

METHODS A narrative review was undertaken of the (published) evidence that has resulted from non-grades-based school-specific selection procedures in the Netherlands and

from recent explorations of the effect of the use of non-grades-based selection criteria on student diversity.

RESULTS The Dutch evidence is grouped into five key themes: the effect of participation in voluntary selection procedures, the assessment of pre-university extracurricular activities, the use of work samples, Dutch experiences with situational judgement tests and the effects of changing circumstances. This is followed by several lessons learned for medical schools that aim to increase their student diversity.

CONCLUSION Over the last 30 years, important steps towards reliable and valid methods for measuring non-academic abilities have been taken. The current paper describes several lessons that can be learned from the steps taken in the Dutch context. The importance of sharing evidence gathered around the globe and building on this evidence to reach our goal of predicting who will be a good doctor is acknowledged.

Medical Education 2018; 52: 86–95
doi: 10.1111/medu.13462



This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Institute of Medical Education Research Rotterdam, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, The Netherlands

Correspondence: K M Stegers-Jager, Institute of Medical Education Research Rotterdam, Erasmus MC, University Medical Centre Rotterdam, Room AE-241, PO Box 2040, 3000 CA Rotterdam, The Netherlands. Tel: 00 31 10 703 8123;
E-mail: k.stegers-jager@erasmusmc.nl

 INTRODUCTION

According to the 1988 Edinburgh Declaration,¹ the aim of medical education is 'to produce doctors who will promote the health of all people'. At that time, that aim was not being realised in many places.¹ Unfortunately, until today, we have still not realised that aim. In 2006, Betancourt described the role medical schools had to play in increasing the proportion of under-represented minorities in the health care workforce in order to eliminate the on-going racial and ethnic disparities.² The promise of improved health care provision, in particular for minority populations, is also one of the arguments for current widening access initiatives.^{3,4} Despite these initiatives, in particular in the USA and the UK, non-traditional students, such as ethnic minority and first-generation university students, are still under-represented in medical schools.^{3,5} Nevertheless, there has been progress, probably because of actions taken by medical schools. One of the actions suggested by the Edinburgh Declaration was to 'Employ selection methods for medical students which go beyond intellectual ability and academic achievement to include evaluation of personal qualities'.

A first rationale behind extending medical school selection procedures with the evaluation of personal qualities is that this will lead to the selection of students who will perform better as a doctor than those who are selected on the basis of academic measures only. However, what personal qualities should one be looking for? The list of desirable personal qualities is likely to be endless; a 2003 review of the literature revealed up to 87 different qualities relevant to the practice of medicine.⁶ So far, attempts to measure personal qualities in selection have been focused on skills such as ethical decision making, communication and collaboration skills.^{4,7,8} Other attempts have been focused on measuring desirable traits in applicants, such as conscientiousness^{9,10} and empathy.¹¹ Recently, there have been calls to also include potential for creativity and innovation as a selection criterion for medical school.¹² Non-academic skills are often measured using contextualised instruments, such as the multiple mini-interview (MMI), consisting of a series of short, structured interviews,^{7,13} or situational judgement tests (SJT), presenting responses to challenging job-related (or medical school-related) situations that applicants have to

judge for their appropriateness.^{14,15} Current evidence shows that attempts to measure non-academic skills using the contextualised MMIs⁷ or SJTs^{4,8} are more effective than attempts to measure desirable traits using self-report questionnaires.^{16,17}

A second rationale for the inclusion of criteria other than grades is to ensure a representative health workforce, as this will lead to improved health care provision, in particular for minority populations.^{3,4} As the use of grades has been shown to introduce a significant bias regarding socio-economic class, especially in the UK,³ the promise of the use of non-grades-based selection is that this will lower adverse impact on non-traditional students. Unfortunately, the introduction of non-grades-based selection criteria has shown mixed results with respect to student diversity.^{3,4,18–20} It was noted that adding measures with reduced adverse impact to existing selection procedures can yield only modest reductions in adverse impacts and sometimes even can have negative effects.²¹

The aims of selecting the best (future) doctors and ensuring a representative health care workforce are further complicated by the following unresolved issues.

The main responsibility of medical schools is to produce competent doctors who will serve societal needs. Therefore, the goal of selection is to identify those who will be successful in medical training and who will ultimately become competent doctors. This in itself already presents a first dilemma for medical schools: should the focus be on selecting the best students or the best doctors? Put differently, which outcomes do we want to predict? This relates to an ongoing debate in selection research: whether or not links to post-academic outcomes (i.e. performance as a doctor) are necessary to justify the use of a particular selection method.²²

If we would like to focus on the best doctors, we are confronted with another dilemma: What is the definition of a good doctor? Do we need one type of doctor or different types? A related difficulty is how to find credible measures of performance in future practice that can be related to the personal qualities assessed during selection. Currently, measures used include supervisor ratings^{8,23} and incidence of remedial action²³ in postgraduate training or results on national licensure examinations,^{7,24,25} with promising outcomes. Nevertheless, it remains extremely difficult to

predict at the time of admission to medical school who will be a good doctor many years later.

A focus on selecting the best students raises additional questions. Who will be the best-performing students depends on the kind of performance you are looking for. Ideally, there is an alignment between admission practices, curriculum (objectives) and assessment practices. This was nicely illustrated by Lievens and colleagues,²⁶ who found that a selection tool measuring interpersonal skills only showed incremental validity over cognitively oriented measures for curricula that included interpersonal courses.

Another issue is whether it would be more effective to select students based on personal qualities or skills or to teach and develop these qualities or skills at medical school. Although evidence shows that at least some qualities are trainable,²⁷ others argue that specific skills and qualities, next to high academic ability, may need to be present from the start.¹⁷ One might argue that the aim would be to select applicants with the highest potential to develop the desirable non-academic skills during medical training. A related question is whether the focus should be on 'selecting in' or on 'selecting out' on the basis of (unsuitable) personal qualities.²⁸

The call for selection procedures that also include the evaluation of personal qualities has led to a divide between so-called cognitive (or academic or intellectual) and non-cognitive (or non-academic or non-intellectual) tests or criteria. However, non-cognitive instruments are very likely to also include cognitive components,²⁹ and cognitive measures like pre-university grade point average (pu-GPA) might also reflect personal qualities such as efficient study strategies and time management.³⁰ Additionally, several studies have shown that non-academic and academic qualities are not independent: success on non-academic criteria enables success on academic criteria and vice versa.^{7,31,32} Hence it might be more appropriate to speak about 'broadened admission criteria',²¹ referring to criteria beyond traditional measures of intellectual ability and pu-GPA.

A first aim of this paper is to describe what can be learned about the use of selection criteria other than grades from over 15 years of Dutch experiences. A second aim is to summarise current knowledge on the issue of adverse impact in relation to non-grades-based selection. Rather than

a systematic review of the literature, I conducted a narrative review guided by my own experience with research into non-grades-based selection and by the evidence that has resulted from more than 15 years of non-grades-based school-specific selection procedures in the Netherlands. The Dutch (published) evidence is grouped into five key themes: the effect of participation in voluntary selection procedures, the assessment of pre-university extracurricular activities, the use of work samples, Dutch experiences with situational judgement tests and the effects of changing circumstances.

NON-GRADES-BASED SELECTION IN THE NETHERLANDS

In the Netherlands, since 2000 until very recently, students were selected for medical school either on the basis of a national lottery system that was weighted for pre-university grades or on medical school-specific selection procedures. The medical schools themselves could decide on their selection criteria, but were not allowed to use pre-university grades. The introduction of these school-specific procedures next to the national lottery system created opportunities for natural experimentation. Additionally, right from the start, medical schools had to consider criteria other than grades. The percentage of students who are admitted via the school-specific procedures has steadily increased from a maximum of 50% in 2000 to currently 100%. Starting from the academic year 2017–2018, the national lottery system was abolished. With this abolishment Dutch medical schools are now permitted to include pre-university grades in their school-specific procedures. The evidence presented below is based on the situation where admission by lottery and by school-specific selection co-existed.

A participation effect?

Until very recently, a unique feature of the Dutch admission system was that medical school applicants could decide whether or not to participate in school-specific selection procedures. Even students who were initially rejected in one of the voluntary school-specific selection procedures still had a chance of admission via the lottery system in the same year. This situation offered the unique possibility to study the effects of the school-specific selection procedures, including the effects of participation.

One of the studies by our group showed a significant lower dropout rate for participants than for non-participants in our two-phase selection procedure.³³ An explanation offered for the difference in dropout was self-selection instigated by the selection procedure. Because our selection procedure required high levels of both the quality and quantity of extracurricular activities, potential applicants may have decided beforehand not to apply because they would not meet these criteria. An alternative explanation is that participation in the selection procedure is associated with a higher motivation to become a medical doctor, reflected in grasping this additional chance to enter medical school.³³ Although this explanation has also been suggested by others,³¹ a recent multi-site study failed to confirm the expected higher motivation for participants than for non-participants in school-specific selection procedures.³⁴ Given the time-consuming nature of most Dutch selection procedures,^{31,35} it was not surprising that participants were found to score higher on conscientiousness than non-participants.³⁶

Other studies, from Denmark and the Netherlands, also revealed this so-called participation effect: non-participants were at higher risk of dropout,³⁷ study delay³¹ and lower grades³¹ than participants. Apparently, in predicting academic performance, *participation* in a time-consuming voluntary selection procedure is more important than *acceptance*.^{31,33,37} Unfortunately, the earlier reported participation effects with respect to study delay and lower grades were not consistently confirmed in a follow-up multi-site study.³⁸ Hence, the conclusion was that participation effects seem to be mediated by institutional differences in curricula and in selection procedures.³⁸

Assessing pre-university extracurricular activities

Several Dutch medical schools, including our own, have operationalised non-academic skills by examining applicants' extracurricular activities during pre-university education (puECAs).^{31,33} This method has the advantage of an increased authenticity, because it reflects students' development over the last couple of years instead of being based on a 'single' test administration. In our school, for example, applicants are assessed on the quality and quantity of their extracurricular activities 2.5 years before application, based on verifiable information provided on the application form. Extracurricular activities include paid and unpaid jobs in health care, experience in

management and organisation, or those that show special talents in domains such as music, science or sports (see Urlings-Strop et al. 2009³⁵ for a more extensive description of our selection procedure).

The first studies by our group showed that students selected by our two-phase selection procedure, including the assessment of puECAs, had a 2.6 times lower relative risk of dropout than students admitted by lottery³⁵ and received significantly higher mean grades on their first five clerkships.³⁹ A Danish study also showed encouraging results in preventing dropout by using puECAs as part of their attributes-based selection procedure.³⁷ Another Dutch study found that their multifaceted selection procedure, including puECAs, was particularly efficient in identifying applicants with suitable 'non-academic' skills, such as professionalism.³¹

A follow-up study from our group focusing on the relative contribution of the different facets of the selection procedure revealed that whereas both the non-academic (i.e. puECAs) and academic (i.e. cognitive tests) selection criteria were associated with a lower chance of dropout during medical school, the better clinical performance of selected students was almost exclusively related to the non-academic selection criteria.³³ In a subsequent study we compared students selected (exclusively) on non-academic criteria and students selected (exclusively) on academic criteria with lottery-admitted students. Contrary to our expectations, students selected on non-academic criteria did not outperform lottery-admitted students in pre-clinical training, whereas those selected on academic criteria did.⁴⁰ This led us to conclude that apparently the use of non-academic selection criteria is not sufficient to select the best academically performing students, probably because a minimum academic level is required to succeed in medical school. It should be noted, however, that this study did not look at differences in clinical performance between the three groups. In our most recent study we found that students selected on the basis of their puECAs persisted in their ECAs during medical school and that this persistent participation was related to better clinical performance, which further supports the inclusion of puECAs in the selection procedure.⁴¹

It is good to note that although puECAs can be seen as a form of personal statements,¹⁷ their assessment, at least in the Dutch medical schools, is different from analysing the content of free

response personal statements, as for example described by Ferguson et al.⁴² The use of a highly structured application form in combination with a highly structured rating form provides a high degree of structure to the raters' judgements. This may positively affect the criterion-related validity, in the same way as job interviews are positively affected by the degree of structure.⁴³ This may explain our positive findings compared with those on personal statements recently summarised in a review by Patterson et al.¹⁷ Consequently, puECAs may be used successfully as a selection tool provided that the judgement process is sufficiently structured.

Work samples

The use of work samples, well known in personnel selection, refers to a situation where a sample of future expected behaviour is taken as a predictor of future success in the job.⁴⁴ The focus on *samples* instead of *signs* (i.e. distinguishable constructs, traits or skills) distinguishes them from traditional selection methods used in medical education.⁴⁵ The reasoning behind this approach is that, following the notion of behavioural consistency,⁴⁴ the more the predictor and criterion are alike, the higher the predictive validity will be.²¹ As early performance is the best predictor of future (medical school) performance,^{46–48} developing an admission test similar to the first course in the university programme sounds promising.

In the context of medical school selection, Lievens and Coetsier⁴⁵ found a correlation of 0.19 for two 'miniaturised' work samples (i.e. a videotaped lecture and reading a medical text) with first-year GPA and concluded that these work samples could significantly better predict first-year GPA than cognitive ability tests. In a more recent study De Visser et al.⁴⁹ showed additional value of their curriculum sample selection procedure compared with pu-GPA, in particular for the lower pu-GPA categories. Other recent Dutch studies also showed positive results for the use of curriculum samples for the selection of undergraduate psychology applicants.^{46,50,51} In all the above studies, applicants were required to watch a videotaped lecture or study introductory domain-specific material independently at home, followed by an examination at the university, just like the reality at university. Positive findings for this efficient method for large groups⁴⁹ are explained by the high representativeness of the sampled behaviour for the educational programme it was designed for.⁵¹ Actually, the selection procedures applied in

Amsterdam, Rotterdam and Groningen also contain 'lecture tests'.^{34,35}

The question of course is, what exactly is measured by sample-based tests, and whether such tests should be considered academic or non-academic measures. De Visser et al.⁴⁹ describe their curriculum sample approach as selection on an academic basis, although it could also be argued that the sample-based approaches described above measure a mixture of cognitive ability, motivation, time spent studying, and tactic knowledge.²¹ At least they measure applicants' performance on authentic tasks that represent behaviour that is required of them during the pre-clinical years of medical school, which might explain the additional value of the tests over pu-GPA. It is important to note that the examples mentioned here are focused on selecting the best students rather than the best doctors.

Situational judgement tests

An example of a non-academic measure that is increasingly used for selection into medical school is the situational judgement test (SJT). As mentioned above, SJTs present applicants with several situations that they may encounter during the job (or at medical school), followed by a number of possible responses to that situation, of which they have to judge the appropriateness. SJTs have consistently been shown to be a reliable and valid method to assess various personal qualities that are important to medical doctors, such as integrity, empathy and teamwork.¹⁷ Thanks to these positive findings, Dutch medical schools have started using (or piloting) SJTs as part of their school-specific selection procedures.

As a first example, the medical school in Groningen used an SJT to measure professionalism in various medical contexts and found that both previous academic experience and a good fit between applicants' vocational interests and SJT scenarios were related to better SJT performance.⁵² As suggested by the researchers, SJTs could hence be used to select appropriate candidates for areas of health care that are in need of more professionals.⁵²

At Erasmus MC Medical School we first pilot-tested an integrity-based SJT developed in the UK,⁵³ which was translated into Dutch. We learned two important lessons from this endeavour. First, one should be cautious in using SJTs developed abroad and tailored to the specific contexts. There were

substantial differences between the original UK scoring key and the scoring key we developed using Dutch experts. Additionally, not all items were considered relevant or appropriate to the Dutch context by our participants. Finally, using the same scoring method as in the UK context revealed a low internal consistency reliability. Following on from these findings we decided to explore 28 (!) alternative scoring methods.⁵⁴ This revealed a second important lesson: the increased use of SJTs in medical school selection must be accompanied by a thorough examination of the scoring method to be used. We found that the applied scoring method has a strong influence on the internal consistency reliability and adverse impact of an SJT score.⁵⁴ Currently, we are in the process of pilot-testing an in-house-developed integrity-based SJT tailored to our own context. Preliminary findings suggest that scoring an SJT based on the ability to recognise what one should not do as opposed to the ability to recognise what one should do in a challenging situation strengthens the convergent validity of that SJT (unpublished results).

Unexpectedly, our intention to include integrity-related aspects in our selection procedure has caused some controversy in the Netherlands. Adversaries doubt whether it is fair to judge young adults on integrity-related issues and propose that these behaviours are teachable. Therefore, it might be better to 'red-flag' applicants with low scores on integrity-related measures and offer them (additional) training instead of rejecting them. Others even go so far as to say that it is unethical to report applicants' scores on integrity-related measures to the client (the medical school), and state that such scores should only be reported to the applicants themselves, who then can decide whether or not to continue with their application.^{55,56} This is related to the remark made by Patterson et al.¹⁷ that personality assessments should not be used as a selection method on their own, but rather as input for selection interviews.

However, our integrity-based SJT is intended to measure knowledge and skills (how one should act) rather than personality traits (how you would act). In our opinion, a certain ability to identify appropriate and inappropriate reactions to challenging situations is essential for our future medical doctors. We also expect applicants passing our SJT to be more responsive to training in professional behaviour during medical education than applicants failing the SJT.

Effect of changing circumstances

So far, studies from the Netherlands have reported conflicting evidence, with some reporting positive effects (i.e. better performance for selected students compared with lottery-admitted students) and some, especially more recent studies, reporting little or no effect. A possible explanation is the increased proportion of students admitted through selection (from 50% to 100%). Moreover, the number of medical schools using school-specific selection procedures increased from only a few to currently all. The effect of increasing the percentage of students admitted by selection on the added value of selection to lottery can be explained using the Taylor-Russell model. This model calculates success ratios (equals the proportion of admitted students who will be successful in medical school and as a doctor) as a function of the base rate, selection ratio and predictive validity of the test(s) used for selection (see also Niessen & Meijer 2016⁵⁷). When the selection ratio (equals percentage of selected applicants) increases without changes in the base rate (equals percentage of suitable applicants), the additional value of selection decreases. Data from the time when all medical students were admitted via lottery show that for all Dutch medical schools the graduation rate was about 85%,⁵⁸ suggesting a high base rate. A main reason for this high base rate is the fact that medical school applicants in the Netherlands form a very homogenous group.⁴⁰ Only students who completed the highest level of secondary school (equals pre-university level) and took the six subjects required for medical school are allowed to apply (less than 10% of all high school students).

Another explanation might be that applicants' behaviour or the rationale behind this behaviour change when admission criteria become known. For example, considering our own context, it might be that, following Higgins' regulatory focus theory,^{59,60} when the requirement of puECA participation for admission became more transparent, applicants may have chosen to participate in puECAs because they felt they 'had to' do this to have a chance to get into medical school, and not because they 'wanted to'. Exploring the regulatory focus of currently admitted students and relations with performance may be an interesting area for further research. The more general lesson from this is that medical schools should be aware of the possible effects of selection procedures on applicants' behaviour. This is especially important as applicants have been

shown to base their choice of a medical school more often on the selection procedure than on the curriculum.⁶¹

ADVERSE IMPACT

As noted previously, the introduction of broader admission criteria has shown mixed results with respect to student diversity.^{3,4,18–20} Nevertheless, recent efforts have led to several lessons for medical schools that aim to increase the diversity of their student population.

First, the effect of using an instrument with reduced or no adverse impact on student diversity depends on the weight given to the instrument and the selection ratio. This was illustrated by Lievens et al.,⁴ who found that the positive effect on the representation of lower SES candidates was stronger when the UKCAT SJT had more weight (i.e. 50%) and the selection ratio was more stringent (i.e. selecting the top 15%). Likewise, we have advocated previously in this journal that medical schools should explore the impacts of different weightings of non-academic and academic criteria on student diversity.³²

Second, different selection criteria should be used concurrently.³² This prevents the diversity-limiting effect of considering personal qualities only once grade requirements have been met.^{20,62} Additionally, to increase diversity it might be better to allow applicants to make up for lower scores on some criteria with high scores on others (compensatory) instead of expecting them to score highly on all selection criteria (non-compensatory).³² Currently, at the Erasmus MC medical school a compensatory system is used. Scores on the three different elements of our selection procedure (a mixture of academic and non-academic criteria) are transformed into z-scores, and applicants are ranked according to their average z-score.

Third, including non-academic criteria in the selection procedures does help to increase the diversity of medical student populations.³² For example, our non-academic selection criterion (i.e. the use of extracurricular activities) did not show an adverse impact with respect to ethnicity or social background.³² However, as advocated previously in this paper, selection procedures should not rely on non-academic criteria alone, but should also include academic criteria.^{33,40}

Fourth, one should be aware of self-selection (i.e. resulting in not applying) in non-traditional applicants. As we have argued previously,³² it cannot be ruled out that self-selection instigated by selection procedures is stronger in applicants from non-traditional backgrounds. Although the percentage of non-Western minority applicants in our study by far exceeded the percentage of non-Western pre-university graduates in the Netherlands (27% versus 8%) and the percentage of first-generation students was higher among applicants than among current students (including those admitted by lottery), we still do not have any data about those who decided not to apply. A recent qualitative study, albeit small, from Amsterdam suggests that selection procedures currently applied in the Netherlands may discourage students without a 'medical network' from applying to medical school, therefore leading to inequalities with respect to socio-demographic background.⁶³ An interesting area of research would be whether certain types of (non-academic) selection instruments are less discouraging for non-traditional applicants, for example because of higher face validity.

Finally, the choice of selection procedure may also influence the gender composition of the student population. Aptitude tests were found to favour men (i.e. male applicants had a higher chance of being admitted based on aptitude tests than female applicants),⁶⁴ whereas instruments measuring non-academic qualities such as SJTs seem to favour female students.⁴ In the Netherlands, where currently two-thirds of the students are female, female students were more likely to be admitted through a multifaceted selection procedure that included reflection assignments and MMI-like interviews³¹ or SJT scenarios,⁵² whereas the use of puECAs as selection criterion did not show gender-related differences.³²

CONCLUSION

Summing up, although there are still multiple issues to be solved regarding the use of selection criteria other than grades, several globally applicable lessons can be learned from the Dutch 'natural experiments'. First, there appears to be a participation effect, in that students who participated in voluntary selection procedures performed better at medical school than non-participants. However, the magnitude of the effect depends on the curriculum and the

selection procedure of the medical school. Second, using pre-university extracurricular activities for selection to medical school seems to lead to lower dropout rates and better pre-clinical and clinical performance, provided that the judgement process is sufficiently structured. Third, the use of curriculum samples is another promising selection method, particularly when the aim is to select those students who will perform best in your curriculum. Fourth, contextualised instruments such as SJTs should really be tailored to the local context and should be accompanied by a careful examination of the scoring method to be used. Fifth, one should be aware of the possibility of controversy when including the assessment of (certain) personal qualities in selection procedures. Sixth, changing circumstances, such as higher selection ratios or applicants' adapting behaviour to meet the selection criteria, may decrease the predictive validity of selection methods. Finally, with respect to adverse impact the main recommendation for medical schools would be to carefully think about a combination and weighting of academic as well as non-academic selection instruments that would fit both the needs of validity and of diversity.^{4,17,32,65}

Issues regarding reliability, validity and fairness of selection procedures have led to recent calls from the UK, the US and the Netherlands to replace medical school selection with lotteries.^{66–68}

However, my opinion is that, as long as we prevent our selection procedures from resembling expensive lotteries⁶⁹ by ensuring their reliability and validity while also taking the diversity amongst applicants into account, it is preferable to consider perceptions of applicants who favour methods that put them 'in control'⁷⁰ as well. Over the last 30 years, important steps towards reliable and valid selection methods for measuring non-academic abilities have been taken, several of them in the Dutch context as described in the current paper. Our task of predicting who will be a good doctor will remain extremely complex and there is much that we still have to understand. While acknowledging this, it is important to share and build on the evidence that is gathered around the globe in order to get a little closer to our goal.

Contributors: the author agrees to be accountable for all aspects of the work in ensuring that questions related to

the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements: the author thanks Axel Themmen and Wendy de Leng for their critical review of the manuscript.

Funding: none.

Conflicts of interest: none.

Ethical approval: no ethical approval was sought for this project.

REFERENCES

- 1 World Federation for Medical Education. The Edinburgh Declaration. *Med Educ* 1988;**22** (5): 481–2.
- 2 Betancourt JR. Eliminating racial and ethnic disparities in health care: what is the role of academic medicine? *Acad Med* 2006;**81**:788–92.
- 3 Cleland J, Dowell J, McLachlan J, Nicholson S, Patterson F. *Identifying best practice in the selection of medical students*. London: General Medical Council 2012. http://www.gmc-uk.org/Identifying_best_practice_in_the_selection_of_medical_students.pdf_51119804.pdf. [Accessed 28 February 2017.]
- 4 Lievens F, Patterson F, Corstjens J, Martin S, Nicholson S. Widening access in selection using situational judgement tests: evidence from the UKCAT. *Med Educ* 2016;**50** (6):624–36.
- 5 Mathers J, Parry J. Why are there so few working-class applicants to medical schools? Learning from the success stories. *Med Educ* 2009;**43** (3):219–28.
- 6 Albanese MA, Snow MH, Skochelak SE, Huggett KN, Farrell PM. Assessing personal qualities in medical school admissions. *Acad Med* 2003;**78**:313–21.
- 7 Eva KW, Reiter HI, Trinh K, Wasi P, Rosenfeld J, Norman GR. Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ* 2009;**43** (8):767–75.
- 8 Lievens F. Adjusting medical school admission: assessing interpersonal skills using situational judgement tests. *Med Educ* 2013;**47** (2):182–9.
- 9 Ferguson E, Semper H, Yates J, Fitzgerald J, A S, James D. The 'dark side' and 'bright side' of personality: when too much conscientiousness and too little anxiety are detrimental with respect to the acquisition of medical knowledge and skill. *PLoS ONE* 2014;**9**:e88606.
- 10 Lievens F, Coetsier P, De Fruyt F, De Maeseneer J. Medical students' personality characteristics and academic performance: a five-factor model perspective. *Med Educ* 2002;**36** (11):1050–6.
- 11 Adam J, Bore M, Childs R, Dunn J, McKendree J, Munro D, Powis D. Predictors of professional behaviour and academic outcomes in a UK medical school: a longitudinal cohort study. *Med Teach* 2015;**37**:868–80.
- 12 Patterson F, Zibarras L. Selecting for creativity and innovation potential: implications for practice in

- healthcare education. *Adv Health Sci Educ* 2017;**22**:417–28.
- 13 Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ* 2004;**38** (3):314–26.
 - 14 Motowidlo SJ, Dunnette MD, Carter GW. An alternative selection procedure: the low-fidelity simulation. *J Appl Psychol* 1990;**75**:640–7.
 - 15 Weekley JA, Ployhart REH, Holtz BC. On the development of situational judgement tests: issues in item development, scaling and scoring. In: Weekley JA, Ployhart RE, eds. *Situational Judgment Tests Theory, Measurement and Application*. New Jersey: Lawrence Erlbaum Associates Inc 2013;157–82.
 - 16 Norman G. Non-cognitive factors in health sciences education: from the clinic floor to the cutting room floor. *Adv Health Sci Educ* 2010;**15**:1–8.
 - 17 Patterson F, Knight A, Dowell J, Nicholson S, Cousans F, Cleland J. How effective are selection methods in medical education? A systematic review *Med Educ* 2016;**50** (1):36–60.
 - 18 James D, Yates J, Nicholson S. Comparison of A level and UKCAT performance in students applying to UK medical and dental schools in 2006: cohort study. *BMJ* 2010;**349**:c478.
 - 19 Puddey IB, Mercer A. Socio-economic predictors of performance in the Undergraduate Medicine and Health Sciences Admission Test (UMAT). *BMC Med Educ* 2013;**13**:155.
 - 20 Reiter HI, Lockyer J, Ziola B, Courneya CA, Eva K. Canadian Multiple Mini-Interview Research A. Should efforts in favor of medical student diversity be focused during admissions or farther upstream? *Acad Med* 2012;**87**:443–8.
 - 21 Niessen S, Meijer R. On the use of broadened admission criteria in higher education. *Perspect Psychol Sci* 2017;**12**:436–48.
 - 22 Lievens F, Sackett PR. The Validity of Interpersonal Skills Assessment Via Situational Judgment Tests for Predicting Academic Success and Job Performance. *J Appl Psychol* 2012;**97**:460–8.
 - 23 Cousans F, Patterson F, Edwards H, Walker K, McLachlan J, Good D. Evaluating the complementary roles of an SJT and academic assessment for entry into clinical practice. *Adv Health Sci Educ* 2017;**22**:401–13.
 - 24 Dore K, Reiter HI, Kreuger S, Norman G. CASPer, an online pre-interview screen for personal/professional characteristics: prediction of national licensure scores. *Adv Health Sci Educ* 2017;**22**:327–36.
 - 25 Eva KW, Reiter HI, Rosenfeld J, Trinh K, Wood TJ, Norman GR. Association between a medical school admission process using the Multiple Mini-interview and National Licensing Examination Scores. *JAMA* 2012;**308**:2233–40.
 - 26 Lievens F, Buyse T, Sackett PR. The operational validity of a video-based situational judgment test for medical college admissions: illustrating the importance of matching predictor and criterion construct domains. *J Appl Psychol* 2005;**90**:442–52.
 - 27 Hook KM, Pfeiffer CA. Impact of a new curriculum on medical students' interpersonal and interviewing skills. *Med Educ* 2007;**41** (2):154–9.
 - 28 Powis D. Selecting medical students: an unresolved challenge. *Med Teach* 2015;**37**:252–60.
 - 29 Prideaux D, Roberts C, Eva K, Centeno A, McCrorie P, McManus C, Patterson F, Powis D, Tekian A, Wilkinson D. Assessment for selection for the health care professions and specialty training: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach* 2011;**33**:215–23.
 - 30 Schripsema NR, van Trigt AM, Borleffs JCC, Cohen-Schotanus J. Underlying factors in medical school admissions. *Med Educ* 2015;**49** (6):639–40.
 - 31 Schripsema NR, van Trigt AM, Borleffs JCC, Cohen-Schotanus J. Selection and study performance: comparing three admission processes within one medical school. *Med Educ* 2014;**48** (12):1201–10.
 - 32 Stegers-Jager KM, Steyerberg EW, Lucieer SM, Themmen APN. Ethnic and social disparities in performance on medical school selection criteria. *Med Educ* 2015;**49** (1):124–33.
 - 33 Urlings-Strop LC, Stegers-Jager KM, Stijnen T, Themmen AP. Academic and non-academic selection criteria in predicting medical school performance. *Med Teach* 2013;**35** (1):497–502.
 - 34 Wouters A, Croiset G, Schripsema NR, Cohen-Schotanus J, Spaai GW, Hulsman RL, Kusurkar RA. A multi-site study on medical school selection, performance, motivation and engagement. *Adv Health Sci Educ* 2017;**22**:447–62.
 - 35 Urlings-Strop LC, Stijnen T, Themmen APN, Splinter TAW. Selection of medical students: a controlled experiment. *Med Educ* 2009;**43** (2):175–83.
 - 36 Schripsema NR, van Trigt AM, van der Wal MA, Cohen-Schotanus J. How Different Medical School Selection Processes Call upon Different Personality Characteristics. *PLoS ONE* 2016;**11**:e0150645.
 - 37 O'Neill L, Hartvigsen J, Wallstedt B, Korsholm L, Eika B. Medical school dropout - testing at admission versus selection by highest grades as predictors. *Med Educ* 2011;**45** (11):1111–20.
 - 38 Schripsema NR, van Trigt AM, Lucieer SM, Wouters A, Croiset G, Themmen AP, Borleffs JC, Cohen-Schotanus J. Participation and selection effects of a voluntary selection process. *Adv Health Sci Educ* 2017;**22** (2):463–76.
 - 39 Urlings-Strop LC, Themmen APN, Stijnen T, Splinter TAW. Selected medical students achieve better than lottery-admitted students during clerkships. *Med Educ* 2011;**46** (10):1032–40.
 - 40 Lucieer SM, Stegers-Jager KM, Rikers R, Themmen APN. Non-cognitive selected students do not outperform lottery-admitted students in the pre-clinical stage of medical school. *Adv Health Sci Educ* 2016;**21**:51–61.

- 41 Urlings-Strop LC, Themmen APN, Stegers-Jager KM. The relationship between extracurricular activities assessed during selection and during medical school and performance. *Adv Health Sci Educ* 2017;**22**:287–98.
- 42 Ferguson E, James D, O'Hehir F, Sanders A. Pilot study of the roles of personality, references, and personal statements in relation to performance over the five years of a medical degree. *BMJ* 2003;**326**:429–31.
- 43 Schmidt FL, Hunter JE. The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychol Bull* 1998;**124**:262–74.
- 44 Wernimont PF, Campbell JP. Signs, samples, and criteria. *J Appl Psychol* 1968;**52**:372–6.
- 45 Lievens F, Coetsier P. Situational tests in student selection: an examination of predictive validity, adverse impact, and construct validity. *Int J Select Assess* 2002;**10**:245–57.
- 46 Visser K, van der Maas H, Engels-Freeke M, Vorst H. Het effect op studiesucces van decentrale selectie middels proefstuderen aan de poort [The effect on study success of student selection through trial studying]. *Tijdschrift voor Hoger Onderwijs* 2012;**30**:161–73.
- 47 Stegers-Jager KM, Themmen APN, Cohen-Schotanus J, Steyerberg EW. Predicting performance: relative importance of students' background and past performance. *Med Educ* 2015;**49** (9):933–45.
- 48 Winston K, Van der Vleuten C, Scherpbier A. Prediction and prevention of failure: an early intervention to assist at-risk medical students. *Med Teach* 2014;**36**:25–31.
- 49 De Visser M, Fluit C, Fransen J, Latijnhouwers M, Cohen-Schotanus J, Laan R. The effect of curriculum sample selection for medical school. *Adv Health Sci Educ* 2017;**22**:43–56.
- 50 Meijer RR, Niessen ASM. A trial studying approach to predict college achievement. *Front Psychol* 2015;**6**:887.
- 51 Niessen ASM, Meijer RR, Tendeiro JN. Predicting performance in higher education using proximal predictors. *PLoS ONE* 2016;**11**:e0153663.
- 52 Schripsema NR, van Trigt AM, Borleffs JCC, Cohen-Schotanus J. Impact of vocational interests, previous academic experience, gender and age on Situational Judgement Test performance. *Adv Health Sci Educ Theory Pract* 2017;**22**:521–32.
- 53 Husbands A, Rodgeron M, Dowell J, Patterson F. Evaluating the validity of an integrity-based situational judgement test for medical school admissions. *BMC Med Educ* 2015;**15**:144.
- 54 De Leng W, Stegers-Jager K, Husbands A, Dowell J, Born M, Themmen A. Scoring method of a Situational Judgment Test: influence on internal consistency reliability, adverse impact and correlation with personality? *Adv Health Sci Educ* 2017;**22**:243–65.
- 55 Hofstee W. De psycholoog als detective? Kanttekeningen bij malingering- en integriteitstests [The psychologist as a detective? Comments on malingering and integrity tests]. *De Psycholoog* 2005;**40**:670–4.
- 56 Hofstee W, Hendriks J, Kiers H. Persoonlijkebeoordeling in de praktijk [Personality assessment in practice]. *De Psycholoog* 2016;**10**:54–8.
- 57 Niessen AS, Meijer RR. Selection of medical students on the basis of non-academic skills: is it worth the trouble? *Clin Med* 2016;**16**:339–42.
- 58 Schmidt HG, Cohen-Schotanus J, Arends LR. Impact of problem-based, active learning on graduation rates for 10 generations of Dutch medical students. *Med Educ* 2009;**43** (3):211–8.
- 59 Higgins ET. Beyond pleasure and pain. *Am Psychol* 1997;**52**:1280–300.
- 60 Higgins ET. Promotion and prevention: regulatory focus as a motivational principle. In: Zanna MP, ed. *Advances in Experimental Social Psychology*. San Diego, CA: Academic Press 1998;1–46.
- 61 Wouters A, Croiset G, Schripsema NR, Cohen-Schotanus J, Spaai G, Hulsman RL, Kusrurkar RA. Student's approaches to medical school choice: relationship with students' characteristics and motivation. *Int J Med Educ* 2017;**8**:217–26.
- 62 Powis D, James D, Ferguson E. Demographic and socio-economic associations with academic attainment (UCAS tariff scores) in applicants to medical school. *Med Educ* 2007;**41** (3):242–9.
- 63 Wouters A, Croiset G, Isik U, Kusrurkar R. Motivation of Dutch high school students from various backgrounds for applying to study medicine: a qualitative study. *BMJ Open* 2017;**7**:e014779.
- 64 Tiffin PA, Dowell JS, McLachlan JC. Widening access to UK medical education for under-represented socioeconomic groups: modelling the impact of the UKCAT in the 2009 cohort. *BMJ* 2012;**344**:e1805.
- 65 Lievens F. Diversity in medical school admission: insights from personnel recruitment and selection. *Med Educ* 2015;**49** (1):11–44.
- 66 Hubbeling D. Lottery for medical school admission. *Med Teach* 2017;**39**:222–3.
- 67 Wouters A. *Effects of medical school selection on the motivation of the student population and applicant pool [dissertation]*. Amsterdam, Netherlands: Vrije Universiteit 2016.
- 68 Conley D. Harvard by Lottery. *Chronicle High Educ*. 2012, 1 April. <http://chronicle.com/article/Harvard-by-Lottery/131322/>. [Accessed 20 June 2017.]
- 69 Norman G. Editorial - The morality of medical school admissions. *Adv Health Sci Educ* 2004;**9**:79–82.
- 70 Niessen ASM, Meijer RR, Tendeiro JN. Applying organizational justice theory to admission into higher education: admission from a student perspective. *Int J Select Assess* 2017;**25**:72–84.

Received 1 March 2017; editorial comments to author 2 May 2017, 26 June 2017; accepted for publication 14 August 2017