

Selectie met gewogen loting

J.A. Smal

Samenvatting

De winst bij selectie op grond van schoolcijfers is afhankelijk van de hoogte van de correlatie tussen de cijfers en het criterium voor succes. Uit twintig jaar onderzoek blijkt dat de correlaties tussen eindcijfers van de middelbare school en resultaten van het eerste studiejaar geneeskunde sterk variëren. De modale hoogte ligt bij $r = +0.35$. Bij een dergelijke correlatie is de winst van gewogen loting in Nederland dat er 4% minder studenten ten onrechte toegelaten zijn dan bij een systeem met pure loting; bij volledige selectie zou deze winst 14% bedragen. Neemt men als criterium voor succes de resultaten in de klinische fase of in de beroepsuitoefening, dan zijn de correlaties met schoolcijfers zo laag dat selectie geen voordeel oplevert.

Inleiding

Twintig jaar geleden is er in Nederland een felle discussie gevoerd over selectie voor het hoger onderwijs.¹ Op dit moment - in 1995 - wordt dezelfde discussie herhaald. De gebruikte argumenten van toen en nu zijn veelal identiek. De psychometrici betogen dat selectie op grond van schoolcijfers geen winst oplevert en bepleiten loting.² Politici geloven in verschillen tussen toppers en subtoppers, en "aan de universiteit behoort geen plaats te zijn voor subtoppers".³ En docenten bekennen zich te schamen als zij hun buitenlandse collega's moeten vertellen dat wij niet selecteren maar loten. Waarom selecteert men in het buitenland wel en doen wij dat niet?⁴ Voor geneeskunde is destijds gekozen voor de 'gewogen' loting, een politiek compromis tussen voor- en tegenstanders van selectie. De hernieuwde discussie

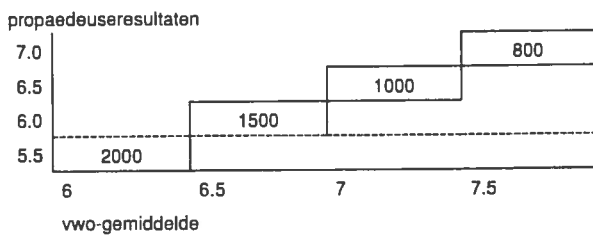
wettigt de vraag of de gewogen loting een aanvaardbare oplossing is voor het selectieprobleem en of nieuwe onderzoeksresultaten aanleiding geven om een andere procedure te bepleiten.

Welke factoren bepalen het nut van selectie?

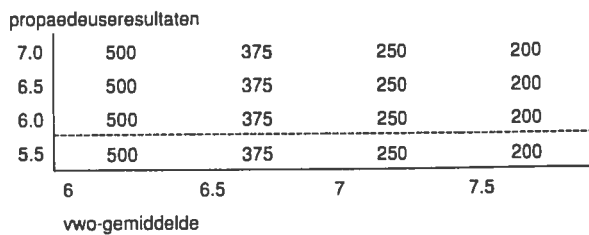
De waarde van selectie is afhankelijk van drie factoren: de gebruikte voorspeller, het criterium voor succes en de correlatie tussen voorspeller en criterium. Bij de toelating tot de medische studie wordt het gemiddelde VWO-cijfer gebruikt als voorspeller en het meest gebruikte criterium voor succes is het propaedeuseresultaat. De correlatie tussen voorspeller en succes kan variëren tussen $r = -1.00$ en $r = +1.00$. Bij een correlatie met een waarde van $r = +1.00$ bestaat er een rechtlijnig verband tussen het gemiddelde VWO-cijfer en het propaedeuseresultaat. In dit geval kan men het studiesucces perfect voorspellen. Een voorbeeld van zo'n rechte lijn is weergegeven in figuur 1.

Daarbij zijn ingevuld de aantallen studenten uit 1994. Er hebben zich toen 5300 studenten aangemeld. Ongeveer 2000 hiervan hadden een cijfer tussen 6.0 en 6.5, ongeveer 1500 tussen 6.5 en 7.0, 1000 tussen 7.0 en 7.5 en 800 hadden een gemiddelde boven de 7.5. De cijfers op de y-as zijn in deze figuur geschat. Daarbij is ervan uitgegaan dat de student in de propaedeuse een half punt lager scoort dan in het VWO en dat ongeveer een derde deel van de studenten de propaedeuse in het eerste jaar niet behaalt.

Uit deze 5300 aanmeldingen mochten er 1700 worden toegelaten. Bij een correlatie van $r = +1.00$ zou selectie op grond van het gemid-



Figuur 1. Verband tussen vwo-gemiddelde en propaedeuse-resultaten bij een correlatie van $r = +1.00$



Figuur 2. Verband tussen vwo-gemiddelde en propaedeuse-resultaten bij een correlatie van $r = 0.00$

delde VWO-cijfer het propaedeuse-rendement verhogen. Immers de 1700 leerlingen met de beste VWO-cijfers zullen ook de beste cijfers krijgen in de propaedeuse. Iedereen zou een 7 of 6.5 halen, dus 100 % behaalt de propaedeuse in een jaar.

Hierbij zijn twee kanttekeningen te maken. Er is allereerst een groep studenten die wel geschikt is, maar niet wordt toegelaten. Dit aspect, de zogenaamde 'onterechte afwijzing' speelt in de discussie tussen voor- en tegenstanders vaak een belangrijke rol. Bij geneeskunde is dit punt van minder belang, omdat door de numerus fixus altijd een groot aantal studenten wordt afgewezen dat in principe ook in staat is de propaedeuse te behalen. Een tweede kanttekening betreft de invulling van de y-as. In dit voorbeeld zou iedereen slagen. Het is echter niet onwaarschijnlijk dat na enige tijd de normen gaan verschuiven, zodat het percentage onvoldoendes na enige tijd toch weer 30% bedraagt. Dit verschijnsel wordt in onderwijskringen wel aangeduid als de 'wet Posthumus'.

Bij een correlatie van $r = 0.00$ bestaat er geen enkel verband tussen voorspeller en criterium: de studenten met lage VWO-cijfers hebben evenveel kans om een goed propaedeuseresultaat te behalen als degenen met een hoog cijfer (figuur 2). Indien men de studenten met de hoogste cijfers selecteert, zal een derde van hen de propaedeuse niet behalen. Maar ook als men een andere groep kiest zijn er evenveel studenten die de propaedeuse niet behalen. In dit geval is loten even goed als selecteren.

De gewogen loting

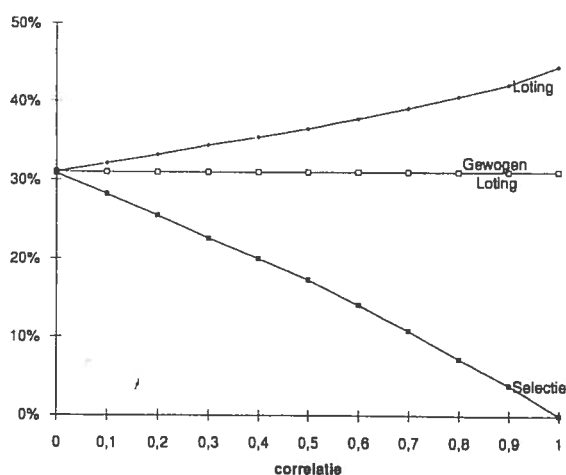
In Nederland wordt voor de toelating tot de studie geneeskunde een combinatie van loting en selectie op schoolcijfers gebruikt in de 'gewogen' loting. Dit wordt bereikt door uit elke cijfergroep evenveel studenten toe te laten. In dit voorbeeld betekent dit 425 studenten uit de groep met het cijfer 6-6.5, 425 uit de groep met het cijfer 6.5-7, 425 uit de groep met het cijfer 7-7.5 en 425 uit de groep met een cijfer hoger dan 7.5. Op deze wijze worden de 1700 toe te laten studenten gekozen. Omdat het aanbod in de cijfergroepen varieert, is ook de kans om ingeloot te worden per cijfergroep verschillend. In de laagste groep heeft de student in dit voorbeeld 21% kans om toegelaten te worden, in de volgende groepen respectievelijk 28%, 43% en 53%. De getallen zijn in dit voorbeeld vereenvoudigd; in de werkelijkheid wordt er ook nog onderscheid gemaakt binnen de groep boven de 7.5.

In figuur 3 zijn de effecten van de drie procedures met elkaar vergeleken bij verschillende correlaties. Daarbij is uitgegaan van de huidige situatie dat 69% van de studenten die via gewogen loting zijn toegelaten de propaedeuse in het eerste jaar behaalt.⁵ Daar de verschillende cijfergroepen in de groep toegelaten studenten anders verdeeld zijn dan in de populatie aangemelde studenten, werd eerst berekend welke resultaten per cijfergroep te verwachten zijn bij verschillende correlatiecoëfficiënten.

ciënten. Hieruit werden de te verwachten resultaten van de gehele populatie afgeleid. Vervolgens werd berekend hoeveel studenten een onvoldoende zouden hebben behaald indien op deze populatie pure loting of selectie op cijfers zou zijn toegepast. De bovenste lijn geeft aan hoeveel studenten ten onrechte worden toegelaten bij pure loting, de middelste lijn geeft het aantal fouten bij gewogen loting en de onderste lijn het aantal fouten bij pure selectie op eindcijfers. Uit de figuur is af te lezen wat bij verschillende correlatiecoëfficiënten de winst is van gewogen loting en van volledige selectie tegenover pure loting.

Uit figuur 3 zijn de volgende conclusies te trekken:

- Bij een correlatie van $r = 0.00$ maakt het niets uit welke selectiemethode gehanteerd wordt;
- Bij hogere correlaties neemt bij selectie het aantal fouten bijna monotoon af en bereikt de waarde nul als de correlatie zijn maximum bereikt ($r = +1.00$);
- Bij lage correlaties is de winst door selectie bescheiden. Bij een correlatie van 0.30 zouden na selectie in plaats van 31% ongeschikten 22% ongeschikten toegelaten zijn;
- Het aantal extra fouten bij loting neemt bijna monotoon toe en groeit tot ongeveer 15% (extra) bij de hoogste correlaties;

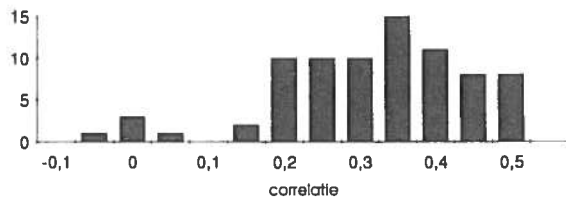


Figuur 3. Percentage onterecht toegelaten bij verschillende correlatiecoëfficiënten

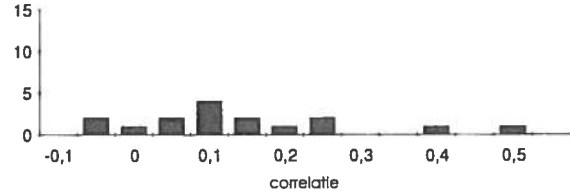
- Het verschil tussen pure loting en gewogen loting is over de hele lijn relatief klein. Wellicht is dit de reden dat er van psychometrische kant weinig protest is gekomen op de keus voor gewogen loting;
- De winst van pure selectie is groter naarmate er minder studenten geschikt blijken: de middelste lijn zou dan veel hoger liggen. Dit betekent dat bij geneeskunde, waar gewoonlijk hoge rendementen behaald worden, selectie niet zoveel winst oplevert.

De hoogte van correlaties in de praktijk

De meerwaarde van selectie op schoolcijfers is duidelijk afhankelijk van de hoogte van de correlatie tussen schoolcijfers en propaedeutische resultaat. De hoogte van de correlatie blijkt in de praktijk echter nogal te variëren. Voorstanders van selectie wijzen op correlaties van $r = +0.50$ die wel eens gevonden zijn, tegenstanders zeggen dat de correlaties gewoonlijk veel lager zijn. Om een betrouwbare indruk te krijgen van de hoogte van correlaties tussen schoolcijfers en resultaten tijdens de medische opleiding is een literatuuronderzoek uitgevoerd. Recente publikaties werden opgespoord in Medline op CD-Rom via de zoektermen "School admission criteria" en "College admission tests". Vanuit de op deze wijze gevonden publikaties werden met de sneeuwbalmethode oudere onderzoeken opgespoord. De zoekperiode is beperkt tot de jaren 1975-1994. Er werd alleen gebruik gemaakt van primaire publikaties waarin correlatie-coëfficiënten vermeld zijn. Op deze wijze werden 14 artikelen getraceerd, die 79 correlatie-coëfficiënten bevatten tussen schooleindcijfers en resultaten in het eerste studiejaar en 16 coëfficiënten van de correlatie tussen schooleindcijfers en prestaties in de klinische jaren van de opleiding. Alle publikaties hadden betrekking op de Verenigde Staten, Groot-Brittannië, Ierland, Canada of Australië.⁶⁻¹⁹ Bij de schoolcijfers werd geen onderscheid gemaakt tussen cijfers voor



Figuur 4. Correlaties tussen middelbare-schoolcijfers en het eerste jaar geneeskunde (n=79)



Figuur 5. Correlaties tussen middelbare-schoolcijfers en de klinische fase (n=16)

afzonderlijke vakken en gemiddelden. Als maat voor studiesucces werden eveneens zowel cijfers voor afzonderlijke vakken als gemiddelden geaccepteerd. Er werd geen onderscheid gemaakt tussen de schooltypen die als voorbereiding op de medical school of universiteit dienden.

Het eerste studiejaar

De gevonden correlaties tussen schoolcijfers en studiesucces in het eerste jaar van de medische opleiding staan weergegeven in figuur 4. Uit deze frequentieverdeling is te zien dat de correlaties meestal rond de $r = 0.35$ liggen. De Nederlandse situatie wijkt hier niet van af.^{20 21} Dit houdt in dat via selectie een foutenreductie van ongeveer 10% mogelijk is. Uit figuur 3 is namelijk af te lezen dat, bij een correlatie van $r = 0.35$, er via selectie 21% onterecht zou worden toegelaten terwijl er bij gewogen loting 31% ten onrechte is toegelaten.

De latere studie jaren en het beroepssucces

Het beeld verandert als men tracht studiesucces in de latere jaren te voorspellen en succes in het beroep. De correlaties met de cijfers in het tweede jaar zijn lager dan die met de cijfers in het eerste studiejaar, en het verband tussen VWO-cijfers en resultaten in de klinische fase is bijna nul (figuur 5).

In enkele grootschalige onderzoeken is het verband tussen schoolcijfers en beroepssucces onderzocht.²² Als maat voor beroepssucces werden tientallen criteria gebruikt. Het resul-

taat was echter volgens de onderzoekers "somewhat disappointing". Van de vele berekende correlaties was 97% niet significant verschillend van $r = 0.00$. De enkele overblijvende moesten beschouwd worden als toevalsafwijkingen.

Conclusie

De hoogte van de correlaties tussen VWO-cijfers en de resultaten in het eerste studiejaar maken de toepassing van selectie op schoolcijfers verdedigbaar. Door de gewogen loting worden 4% minder studenten ten onrechte toegelaten dan het geval zou zijn bij pure loting. Door volledig te selecteren op schoolcijfers zouden nog eens 10% minder studenten ten onrechte worden toegelaten. Kijkt men echter enkele jaren verder, naar de klinische fase bijvoorbeeld, dan blijkt selectie op grond van schoolcijfers niets te voorspellen van het succes in de opleiding. Selectie op schoolcijfers levert derhalve op de langere termijn gezien niets op. Blijft over de mogelijkheid om een andere voorspeller te kiezen, bijvoorbeeld bètavakken, interviews, persoonlijkheidsvragenlijsten of psychologische tests. Ook hier ontmoet men voor- en tegenstanders. De voorstanders wijzen op enkele hoge correlaties, de tegenstanders op lage correlaties.^{23 24} Zoals bij de discussie over de waarde van schoolcijfers als selectiemiddel, is het van belang niet af te gaan op enkele toevallige correlaties, maar de uitkomsten van zoveel mogelijk onderzoek te combineren.

Literatuur

1. Selectiebundel: samenvatting van het rapport van prof. dr. A.D. de Groot "Selectie voor en in het hoger onderwijs" met perscommentaren. Utrecht: Afdeling Public Relations RU, 1973.
2. Drenth PJD. In Nederland is selectie onmogelijk. NRC Handelsblad 1995 maart 30; sectie Wetenschap & Onderwijs: 1-2.
3. Bolkestein F. Selectie is elitair en rechtvaardig. NRC Handelsblad 1995 maart 30; sectie Wetenschap & Onderwijs: 3.
4. Eykhoff P. Wat elders kan, kan hier ook. NRC Handelsblad 1995 maart 30; sectie Wetenschap & Onderwijs: 2.
5. Vereniging Samenwerkende Nederlandse Universiteiten. Visitatierapport geneeskunde en gezondheidswetenschappen. Utrecht: VSNU, 1992.
6. Mawhinney BS. The value of ordinary and advanced level British school-leaving examination results in predicting medical students' academic performance. Med Educ 1976; 10: 87-90.
7. McGuire FL. Fifteen years of predicting medical school performance. J Med Educ 1977; 52: 416-7.
8. Irwin WG, Bamber JH. Characteristics of senior medical students at Belfast. Med Educ 1978; 12: 117-23.
9. Murden R, Galloway GM, Reid JC, Colwill JM. Academic and personal predictors of clinical success in medical school. J Med Educ 1978; 53: 711-9.
10. Lennox B, Dick I, Graham B. Comparison of admission criteria and medical school programs: preliminary results from a simple computer system. Med Educ 1979; 13: 374-9.
11. Cullen TJ, Dohner CW, Peckham PD, Samson WE, Schwarz MR. Predicting first-quarter test scores from the new Medical College Admission Test. J Med Educ 1980; 55: 393-8.
12. Golmon ME, Berry CA. Comparative predictive validity of the new MCAT using different admission criteria. J Med Educ 1981; 56: 981-6.
13. McManus IC. A-level grades and medical school admission. BMJ 1982; 284: 1654-6.
14. Lipton A, Huxham GJ, Hamilton D. Predictors of success in a cohort of medical students. Med Educ 1984; 18: 203-10.
15. Jones RF, Thomae-Forgues M. Validity of the MCAT in predicting performance in the first two years of medical school. J Med Educ 1984; 59: 455-64.
16. Johnson DG, Lloyd SM, Jones RF, Anderson J. Predicting academic performance at a predominantly black school. J Med Educ 1986; 61: 629-39.
17. Nowacek GA, Pullen E, Short J, Blummer HN. Validity of MCAT scores as predictors of preclinical grades and NBME Part I examination scores. J Med Educ 1987; 62: 989-91.
18. Weiss M, Lotan I, Kedar H, Ben-Shakir G. Selecting candidates for a medical school: an evaluation of a selection model based on cognitive and personality predictors. Med Educ 1988; 22: 492-7.
19. Montague W, Odds FC. Academic selection criteria and subsequent performance. Med Educ 1990; 24: 151-7.
20. Baneke JJ. Studiesucces, persoonlijkheid en stress bij geneeskunde studenten. Proefschrift. Amsterdam: Universiteit van Amsterdam, 1987.
21. Jansen, E. De kracht van de vernieuwde tweede fase VWO. Onderzoek van Onderwijs 1995 (feb); 9-11.
22. Taylor CW, Albo D. Measuring and predicting the performances of practicing physicians: an overview of two decades of research at the university of Utah. Academic Medicine 1993; 68: 65-7.
23. Powis DA. Selecting medical students. Med Educ 1994; 28: 441-69.
24. Smal JA, Gerritsma JGM. Selectie. In: Metz JCM, Scherpier AJJA, Van der Vleuten CPM, eds. Medisch onderwijs in de praktijk. Assen: Van Gorcum, 1995: 259-64.

DE AUTEUR

J.A. Smal is hoofd Stafafdeling Onderwijs en Onderzoek van de faculteit Geneeskunde Universiteit Utrecht.

Correspondentie-adres:

J.A. Smal, Afdeling Onderwijs en Onderzoek Geneeskunde, Postbus 80030, 3508 TA Utrecht.