

De kennis van de Nederlandse basisartsen gemeten met de Maastrichtse voortgangstoets

P. A. W. VAN HESSEN, G. M. VERWIJNEN EN T. J. IMBOS

INLEIDING

De Maastrichtse voortgangstoets (MVT) is regelmatig afgenomen bij anderen dan Maastrichtse studenten. De resultaten tonen een vrij constant beeld: studenten van diverse faculteiten scoren gemiddeld hoger naarmate zij verder in het curriculum gevorderd zijn, huisartsen in opleiding en andere onlangs afgestudeerden, samen aangeduid als 'referenten', scoren meestal iets hoger dan studenten in hun laatste studiejaar, en de verschillen tussen faculteiten zijn gering.¹⁻⁶

Begin 1988 stelde de medische faculteit van de Rijksuniversiteit Limburg (RL) middelen ter beschikking om opnieuw referenten te werven. Men wilde beschikken over recente referentenscores als referentiepunten voor de scores van, vooral 6e-jaars, studenten. Pas afgestudeerde artsen (179) van andere universiteiten namen deel als referent aan de MVT van mei 1988. Overeenkomstig gemaakte afspraken ontvingen 51 van hen hiervoor f 100,- en kregen de overige 128 geen vergoeding. De onbetaalden scoorden iets beter.⁷ Zoals verwacht, was de animo om deel te nemen groter bij betaling.

Scores van deze referenten leken veel op die van de RL-6e-jaars. In het verlengde hiervan ontstond de vraag of scores op de MVT van basisartsen van de RL eveneens veel overeenkomst vertonen met de scores van basisartsen van elders. Er zou daarbij ook gekeken dienen te worden in hoeverre de MVT relevant is voor basisartsen. De afname van de toets van mei 1988 werd beschouwd als oriënterend voor het onderzoek van deze vraag bij de toets van mei 1989.

METHODE

De Maastrichtse voortgangstoets. Deze toets is ontwikkeld in het kader van het 'probleemgeoriënteerde' karakter van het Maastrichtse curriculum.⁸⁻¹¹ Hij bestaat uit circa 235 vragen uit het gehele domein van medische kennis op basisartsniveau. Studenten van alle studiejaren maken viermaal per jaar een MVT. Naarmate hun studie vordert, kunnen en moeten zij meer vragen correct beantwoorden. Vragen worden op grond van vakgroep-herkomst ingedeeld in drie clusters: basis-, klinische en gedragswetenschappelijke vakken.

Een vraag moet relevante kennis voor een basisarts betreffen. Hierover beslist uiteindelijk de vakgroep die de vraag gemaakt heeft. Uitdrukkelijk wordt dus niet het criterium gehanteerd dat het onderwerp in het curriculum aan bod moet zijn gekomen.¹² De MVT wordt daarom 'curriculum-onafhankelijk' genoemd.

De vragen bestaan uit stellingen die beantwoord kunnen worden met 'juist', 'onjuist' of een vraagteken. De score van een student is het aantal correcte minus het aantal incorrecte antwoorden, uitgedrukt als proportie van het totale aantal vragen. Dit is de zogenaamde goed-minus-foutscore. Vraagtekens tellen dus niet mee. De scores op MVT'en zijn een belangrijke informatiebron bij het nemen van beslissingen over de voortgang van studenten.

De gemiddelde goed-minus-foutscore van de 6e-jaars studenten ligt al jaren rond de 40%. Volgens sommigen betekent dit dat studenten te weinig weten, volgens anderen dat de vragen toch te moeilijk of te weinig relevant zijn. Het laatste woord hierover is nog niet gezegd.

Voor dit onderzoek bij basisartsen werd gebruik gemaakt van de reguliere MVT van mei 1989, bestaande uit 227 vragen. Ter vergelijking worden soms ook resultaten van studenten, vooral 6e-jaars, op deze toets vermeld.

Deelnemers. In het voorjaar van 1989 werden de 1810 artsen aangeschreven, die waren geregistreerd bij het ministerie van WVC als afgestudeerd (of anderszins erkend) in 1988, en van wie een adres bekend was. Hieronder waren 113 afgestudeerden van de RL. Deze 1810 artsen werd verzocht thuis de MVT te maken. In het oriënterend onderzoek was de respons bij betaling ruim 50% geweest, zonder betaling bijna 10%. Teneinde voldoende RL-artsen te kunnen werven, werd hun een vergoeding van f100,- in het vooruitzicht gesteld; gegeven het beperkte budget gebeurde dat niet bij de overige artsen (in het volgende 'niet-RL-artsen' genoemd). Aan allen werd uitgebreide 'feedback' beloofd. Er namen 375 artsen deel: 67 RL-artsen (50%) en 308 niet-RL-artsen (18%). Onder deze laatsten varieerde het aantal per universiteit van 33 tot 60; 2 artsen waren op andere wijze erkend dan door een regulier artsexamen aan een Nederlandse universiteit.

De deelnemers was gevraagd op te geven wat hun werkzaamheden op dat moment waren. Hierin bleek de deelnemersgroep in grote mate overeen te komen met wat uit eerdere onderzoeken bekend is over de werkzaamheden van artsen 0,5 tot 1,5 jaar na afstuderen.^{13 14}

Rijksuniversiteit Limburg, Faculteit der Geneeskunde, Project Evaluatie Studieresultaten, Postbus 616, 6200 MD Maastricht.

P. A. W. van Hessen (thans: GGD Stadsgewest Breda); G. M. Verwijnen, huisarts; dr. T. J. Imbos, psychometricus.

Correspondentie-adres: G. M. Verwijnen.

Om een indruk te krijgen van de representativiteit van de deelnemende RL-artsen wat betreft hun kennisniveau, werden hun scores op 2 eerdere MVT'en, die zij als 6e-jaars student gemaakt hadden, vergeleken met de gemiddelde waarden van deze MVT'en. Ruim 60% van de artsen had destijds aan de MVT'en deelgenomen, en hun gemiddelde score was beide keren gelijk aan het jaargemiddelde bij afronding op hele procenten. Dat het deelnemerspercentage niet 100 was – wat men zou verwachten bij een verplichte toets – komt door de examenregels en de organisatie in het laatste studiejaar: studenten lopen vaak nog co-assistentenschappen of keuzestages, terwijl zij al eerder aan de vereiste deelname aan de MVT hebben voldaan.

Variabelen. Bijna alle artsen vulden relevantie-oordelen in. Het verzoek was per vraag aan te geven of men deze kennis relevant of irrelevant vond voor een basisarts, of er neutraal tegenover stond.

De interne consistentie van de toets voor een groep deelnemers werd beschreven met behulp van Cronbachs α , een maat voor betrouwbaarheid.¹⁵ Kennisniveaus werden vergeleken aan de hand van goed-minus-foutscores op de totale MVT en per vakkencluster. Omdat deze scores per groep redelijk voldeden aan de eisen van een normale verdeling, konden de verschillen getoetst worden met een t-toets.

Analyse van de aard van de kennis gebeurde met zogenaamde k-waarden. Hierbij was de k-waarde van een vraag in een groep het percentage mensen uit die groep dat de vraag correct beantwoordde minus het percentage dat een incorrect antwoord gaf. Naarmate een k-waarde hoger was, was in de betreffende groep de correcte kennis over het in de vraag getoetste groter. K-waarden in 2 groepen werden beschouwd als bivariate gegevens aangaande toetsvragen, en hun samenhang werd geschat door middel van de berekening van de

lineaire correlatiecoëfficiënt. Om na te gaan of het berekenen van correlaties wel geoorloofd was met een lineair model, werden eerst de regressiecoëfficiënten van de bijbehorende lineaire regressievergelijkingen getoetst op significantie.

RESULTATEN

Relevantie en betrouwbaarheid. In tabel 1 is het oordeel over relevantie weergegeven. De vragen van de klinische vakgroepen achtte men het meest relevant. Steeds was het oordeel van RL-artsen positiever dan van niet-RL-artsen. Per vraag liepen de percentages overigens zeer uiteen.

Voor de totale goed-minus-foutscore op de MVT gold bij de RL-artsen $\alpha = 0,73$ en bij de niet-RL-artsen $\alpha = 0,78$. De betrouwbaarheid bereikte daarmee niet de in de literatuur genoemde wenselijke grens van 0,80, maar was in vergelijking met de gangbare resultaten in de toetspraktijk alleszins acceptabel te noemen.¹⁶ In de groep 6e-jaars was α eveneens gelijk aan 0,78.

Kennisniveau. In tabel 2 staan de gemiddelde goed-minus-foutscores. De niet-RL-artsen scoren hoger dan de RL-artsen. Alleen bij de basisvakken was het verschil, dat daar 4,8% bedroeg, statistisch significant ($p \leq 0,05$). De RL-artsen verschilden weinig van de 6e-jaars studenten. Overigens 'groeiden' de studentenscores op deze MVT met gemiddeld bijna 6% per jaar van 10,3% in jaargroep 1 tot 39,8% in jaargroep 6.

De standaarddeviaties waren groot: 9,1 en 10,6 voor de totale MVT. Variatie in scores was dus groot. Bij vergelijking met de studenten bleek dat 12% van de artsen lager scoorde dan het gemiddelde in jaargroep 4: 29,9%.

Aard van de kennis. Correlaties tussen k-waarden zijn afgebeeld in tabel 3. Een groep afgestudeerden van een bepaalde universiteit werd steeds vergeleken met de

TABEL 1. Gemiddeld percentage van basisartsen afkomstig van de Rijksuniversiteit Limburg (RL) en van niet-RL-basisartsen dat een vraag uit de Maastrichtse voortgangstoets als 'relevant', 'irrelevant' of neutraal beoordeelde

vakkencluster	relevant		irrelevant		neutraal	
	RL-artsen	niet-RL-artsen	RL-artsen	niet-RL-artsen	RL-artsen	niet-RL-artsen
basisvakken (88 vragen)	80	70	12	18	9	12
klinische vakken (104 vragen)	86	82	8	9	7	9
gedragwetenschappelijke vakken (35 vragen)	72	63	15	19	13	18
totaal (227 vragen)	81	74	10	14	9	12

TABEL 2. Gemiddelde goed-minus-foutscores van basisartsen afkomstig van de Rijksuniversiteit Limburg (RL) en van niet-RL-basisartsen op vragen uit de Maastrichtse voortgangstoets (tussen haakjes de standaarddeviatie)

groep	basisvakken (88 vragen)	klinische vakken (104 vragen)	gedragwetenschappelijke vakken (35 vragen)	totaal (227 vragen)
6e-jaars studenten (n = 117)	39,0 (12,6)	40,5 (11,0)	39,9 (13,8)	39,8 (9,6)
RL-artsen (n = 67)	38,1 (11,3)	41,4 (11,0)	37,9 (17,1)	39,6 (9,1)
niet-RL-artsen (n = 308)	42,9 (13,8)	42,5 (10,5)	38,1 (13,3)	42,0 (10,0)
p-waarde*	0,009	0,448	0,917	0,074

* Toetsing van het verschil tussen RL- en niet-RL-artsen (2-zijdige t-toets).

TABEL 3. De lineaire correlaties tussen de gemiddelde goed-minus-foutscores per vraag (zogenaamde k-waarden)* van basisartsen afkomstig van een bepaalde universiteit ten opzichte van de (overige) niet van de universiteit Limburg afkomstige basisartsen

universiteit van afstuderen	aantal deelnemers	correlatiecoëfficiënt†
Groningen	36	0,90
Nijmegen	51	0,92
Utrecht	38	0,93
Amsterdam (UvA)	33	0,88
Amsterdam (VU)	60	0,94
Leiden	46	0,94
Rotterdam	42	0,91
Maastricht	67	0,91

* De k-waarde van een vraag in een groep is het percentage mensen uit die groep dat de vraag correct beantwoordde minus het percentage dat een incorrect antwoord gaf.

† Correlatiecoëfficiënt bij correlatie met (overige) niet-RL-artsen; voor de niet-RL-artsen van een bepaalde universiteit werd de correlatie berekend met de overige niet-RL-artsen, voor de RL-artsen met de totale groep niet-RL-artsen.

overige niet-RL-artsen; de 67 RL-artsen dus met alle 308 niet-RL-artsen. De lineaire regressie-analyses die bij de correlaties behoren, leverden steeds regressie-coëfficiënten op die significant waren ($p \leq 0,05$). De correlatiecoëfficiënten varieerden van 0,88-0,94. De waarde 0,91 voor de RL-artsen was hierbij een gemiddelde waarde.

Deze correlaties lagen in dezelfde orde van grootte als die tussen opeenvolgende jaargroepen studenten bij deze toets (jaargroep 1 en 2, 2 en 3 etcetera): respectievelijk 0,81, 0,89, 0,88, 0,90 en 0,91.

BESCHOUWING

Er moet enig voorbehoud gemaakt worden bij de vergelijkbaarheid van beide groepen artsen (RL-artsen en niet-RL-artsen). De respons van 50% bij RL-artsen en van 18% bij niet-RL-artsen is niet hoog, en berustte op zelfselectie; selectiebias is mogelijk. De RL-artsen werden bovendien betaald voor hun deelname, de niet-RL-artsen niet. De MVT werd thuis gemaakt onder niet gecontroleerde omstandigheden. Of men informatiebronnen raadpleegde, en of hierin verschil bestond tussen RL- en niet-RL-artsen, is niet bekend. Wel vormen de deelnemende artsen wat betreft hun werkzaamheden een redelijke afspiegeling van de populatie van gemiddeld ongeveer een jaar afgestudeerde artsen. De RL-artsen waren ook qua kennisniveau een representatieve groep gezien hun eerdere prestaties op MVT'en. In het oriënterend onderzoek van een jaar eerder bleken weinig verschillen in diverse variabelen te bestaan tussen wel en niet voor deelname betaalde artsen. Dat de artsen op grote schaal informatiebronnen geraadpleegd hebben lijkt ons niet aannemelijk. Uit ervaringen met het opstellen van de MVT is bekend dat dit raadplegen veel tijd kost. De artsen hadden geen belang bij een goed resultaat, alleen zichzelf konden hun individuele resultaten inzien. Bovendien 'gedraagt' de toets zich bij artsen en studenten vergelijkbaar. Dat betreft dan vooral standaarddeviaties, betrouwbaarheden en correlaties van

k-waarden met andere groepen. Wanneer artsen informatiebronnen hadden geraadpleegd was daarin minder overeenkomst te verwachten geweest. De mogelijkheid van het raadplegen van bronnen kan echter niet worden uitgesloten.

Of de MVT een valide toetsinstrument is met betrekking tot kennis van basisartsen, is nooit met zekerheid aan te tonen.¹⁷ Bekend is dat de score van huisartsenin-opleiding op de MVT stijgt gedurende hun beroepsopleiding.² Het afnemen van een deel van een MVT in Leiden en Groningen laat een gelijkmatige stijging van scores zien over de jaren heen, met wat hogere scores voor de referenten.¹ Ondersteuning geeft verder de procedure van samenstelling van de MVT. Hierbij geldt als criterium relevantie voor de basisarts, onafhankelijk van behandeling in het curriculum.¹² De acceptabele toetsbetrouwbaarheid en overwegend positieve relevantie-oordelen bij dit onderzoek ondersteunen de validiteit van de MVT. Dat de RL-artsen de relevantie hoger inschatten dan niet-RL-artsen kan liggen aan hun vertrouwdheid met de MVT als toets voor kennis.

De RL-artsen scoorden lager op de MVT dan de niet-RL-artsen. Dit zou gedeeltelijk verklaard kunnen worden door het al dan niet betaald worden. Ook in het oriënterend onderzoek scoorden de onbetaalde artsen gemiddeld wat hoger dan de betaalde, al was dit verschil niet statistisch significant.⁷ Daarnaast leek men in Maastricht minder te leren van de basisvakken (zie tabel 2). Deze bevinding is niet in overeenstemming met eerder gevonden resultaten bij 8 referentengroepen.⁷ Bij het klinische en gedragswetenschappelijke cluster waren de verschillen kleiner. Voor de totale MVT was er geen significant verschil op $p \leq 0,05$ -niveau. Deze bevinding sluit aan bij een overzichtsartikel uit 1987 over probleemgeoriënteerde versus conventionele curricula in diverse landen. Daarin wordt geconcludeerd dat uit meerdere onderzoeken niet of nauwelijks verschillen in kennis blijken tussen de betreffende studenten.¹⁸

In beide groepen artsen was de variatie in scores groot; dit gold overigens ook voor studenten. Van de artsen scoorde 12% onder het 4e-jaars niveau. Uit de Maastrichtse ervaring is bekend dat er intra-individuele variatie bestaat, hetgeen wil zeggen dat dezelfde student in vergelijking met jaargenoten de ene keer hoger scoort op een MVT dan de andere keer. In het algemeen is die variatie echter gering. Als de MVT een redelijk beeld zou geven van het kennisniveau van artsen, zijn er dus wellicht pas afgestudeerde artsen met een kennisniveau onder dat van een gemiddelde 4e-jaars RL-student.

Bij de studenten worden hoge correlaties tussen de k-waarden in opeenvolgende jaargroepen gevonden. Er bestond dus een grote samenhang in de aard van de kennis tussen jaargroepen. Dat is niet verwonderlijk binnen één medisch curriculum. Er blijkt echter een even grote samenhang te bestaan tussen groepen artsen onderling. De voorspelbaarheid van de scores per vraag bij de artsen van 1 universiteit op grond van de scores van de overige niet-RL-artsen was hoog. De RL-artsen vormden geen uitzondering. Dit duidt op een grote mate van 'curriculum-onafhankelijkheid' in de aard van de medi-

sche kennis zoals die door de MVT gemeten wordt. Zou dit immers niet zo zijn, en zou er bijvoorbeeld een typisch Maastrichts antwoordpatroon op de MVT bestaan, dan zou de voorspelbaarheid van dat patroon laag zijn.

De MVT kan nog steeds typisch Maastrichtse blinde vlekken hebben, en er kan ook zoiets als typisch Maastrichtse kennis bestaan. Maar als een groot deel van het domein van medische kennis bij basisartsen gedekt wordt door de MVT, dan lijkt men op de verschillende universiteiten, waaronder de RL, voor een groot deel hetzelfde (niet) te leren.

Al in 1984 leidde een dergelijke constatering tot het advies interfacultaire samenwerking op toetsgebied te stimuleren.¹ In oktober 1990 is dit weer gesuggereerd tijdens het symposium: 'Basisarts: bevoegd, maar waar toe bekwaam?'¹⁹ Uitwisselen van goed, en dus arbeidsintensief, toetsmateriaal via een landelijke opgavenbank zou efficiëntie en kwaliteit bevorderen. Bovendien kan het toetsmateriaal gebruikt worden voor nader onderzoek naar overeenkomsten en verschillen tussen studenten en (of) artsen van de diverse universiteiten.

LITERATUUR

- 1 Bender W, Cohen-Schotanus J, Imbos T, Versfelt WA, Verwijnen M. Medische kennis bij studenten uit verschillende medische faculteiten: van hetzelfde laken een pak? Ned Tijdschr Geneesk 1984; 128: 917-21.
- 2 Bulte JA, Verwijnen GM, Tielens VCL, Leeuwen YD van. Kennis bij huisartsen in opleiding. Med Contact 1988; 46: 1426-8.
- 3 Verwijnen GM, Vleuten CPM van der, Imbos Tj. A comparison of an innovative medical school with traditional schools: an analysis in the cognitive domain. In: Khattab T, Schmidt HG, Nooman Z, Ezzat E, eds. Innovation in medical education: an evaluation of its present status. New York: Springer, 1989.
- 4 Cate ThJ ten. De algemene medische kennistoets: resultaten. NPOG-memo 86-1. Amsterdam: Bureau Subfaculteit Geneeskunde, Universiteit van Amsterdam, 1986.

- 5 Bosschert D. Resultaten afname voortgangstoets. Leiden: Sectie Onderwijsontwikkeling, Faculteit der Geneeskunde, Rijksuniversiteit, 1987.
- 6 Bosschert D. Resultaten voortgangstoets 30 juni 1988. Leiden: Sectie Onderwijsontwikkeling, Faculteit der Geneeskunde, Rijksuniversiteit, 1988.
- 7 Hessen PAW van. Referenten bij de voortgangstoets van mei 1988. PES-publicatie 89-06. Maastricht: Faculteit der Geneeskunde, Rijksuniversiteit, 1989.
- 8 Wijnen WHFW, Vleuten CPM van der. Toetsing: hordenloop of voortgangscontrole? Universiteit & Hogeschool 1985; 31: 270-9.
- 9 Verwijnen GM, Imbos Tj, Snellen H, et al. The evaluation system of the medical school of Maastricht. Assessment and Evaluation in Higher Education 1982; 3: 225-44.
- 10 Imbos Tj. Het gebruik van einddoeltoetsen bij aanvang van de studie. Maastricht, 1989. Proefschrift.
- 11 Vleuten CPM van der. Naar een rationeel systeem voor toetsing van studieprestaties in probleemgestuurd medisch onderwijs. Amsterdam, 1989. Proefschrift.
- 12 Hessen PAW van. Toetsen getoetst: Het beoordelen van toetsvragen in Maastricht. Bulletin Medisch Onderwijs 1989; 8: 100-5.
- 13 Driel ML van, Kappelhorst J. Basisarts en werk. Utrecht: Instituut voor Sociale Geneeskunde, 1984.
- 14 Commissie werkgelegenheid Landelijke Belangenvereniging voor Basisartsen. Basisarts, en dan ...? Amsterdam: Wetenschapswinkel Vrije Universiteit, 1986.
- 15 Cronbach LJ. Essentials of psychological testing. New York: Harper & Row, 1970.
- 16 Swanson D, Norcini J, Grosso L. Assessment of clinical competence: written and computer-based simulations. Assessment and Evaluation in Higher Education 1987; 12: 220-46.
- 17 Hofstee WKB. Begripsvalidatie van studietoetsen: een aanbeveling. Nederlands Tijdschrift voor Psychologie 1971; 26: 491-500.
- 18 Schmidt HG, Dauphinee WD, Patel VL. Comparing the effects of problem-based and conventional curricula in an international sample. J Med Educ 1987; 62: 305-15.
- 19 Hoedemaker PhJ, Jansen FAJ. Landelijk artsexamen: toetssteen of steen des aanstoots. Bulletin Medisch Onderwijs 1990; 9: 164-7.

Aanvaard op 30 mei 1991

Referaten

Inwendige geneeskunde

Hoge mortaliteit van patiënten met diabetes mellitus die met diuretica worden behandeld

In een retrospectief onderzoek is door Warram et al. nagegaan of de hoge mortaliteit van patiënten met diabetes mellitus die behandeld worden wegens een verhoogde bloeddruk, het gevolg is van de ziekte zelf en de bijkomende risicofactoren of geweten kan worden aan de nadelige gevolgen van de bloeddrukverlagende middelen.¹ Het onderzoek betrof 759 patiënten met diabetes mellitus en een ernstige retinopathie. De gegevens van deze patiënten waren bekend, omdat allen hadden deelgenomen aan een onderzoek waarbij de werkzaamheid van laserbehandeling werd onderzocht bij de behandeling van retinopathie. De patiënten waren blank en hun leeftijd was 35 tot 69 jaar. De meesten onder hen werden behandeld met insuline. Bij alle deelnemers was, bij het eerste onderzoek, het creatininegehalte in het bloed normaal. Bij de verdere analyses werden de patiënten in 5 groepen onderverdeeld: patiënten met een normale diastolische bloeddruk (< 90 mmHg); patiënten met een verhoogde bloeddruk die daar niet voor werden

behandeld; patiënten met hypertensie die behandeld werden met diuretica; patiënten die wegens een verhoogde bloeddruk werden behandeld met andersoortige antihypertensiva en ten slotte patiënten bij wie de hypertensie behandeld werd met de combinatie van diuretica en andere bloeddrukverlagende middelen. De gemiddelde controleduur was 4,5 jaar. Tijdens deze periode overleden 139 patiënten, vooral als gevolg van hart- en vaatziekten. De mortaliteit als gevolg van cardiovasculaire aandoeningen was bij patiënten met een niet behandelde hypertensie lager dan die van patiënten die, wegens de verhoogde bloeddruk, medicamenteus werden behandeld. Deze hoge mortaliteit werd vooral gevonden bij patiënten die uitsluitend met diuretica werden behandeld (terwijl toch bij deze groep de bloeddruk het meeste was gedaald tijdens de therapie). Bij multivariate analyse bleek het gebruik van diuretica een onafhankelijke risicofactor te zijn. Wanneer rekening gehouden werd met verschillende risicofactoren kon worden berekend dat de mortaliteit als gevolg van cardiovasculaire ziekten 3,8 maal zo groot was bij diabetespatiënten die met diuretica werden behandeld in vergelijking met diabetespatiënten met een niet behandelde hypertensie ($p < 0,001$).