

MEDISCHE KENNIS BIJ (BEGINNENDE) CO-ASSISTENTEN

TUSSEN SERVET EN TAFELLAKEN?

J.C.M. METZ
J.A. BULTE

Sinds februari 1987 volgen studenten

in Nijmegen, alvorens aan hun klinische co-assistentschappen te beginnen, het zogenaamde Algemeen Co-assistentschap. Dit ALCO-schap is primair een systematische training in klinische vaardigheden. Het beschikken over een zekere hoeveelheid parate kennis, wordt als voorwaarde gezien voor het goed kunnen uitvoeren van al deze vaardigheden. Op de eerste dag van het ALCO-schap wordt daarom aan de studenten een "Intreetoets" voorgelegd. Dit is een kennistoets op het niveau van beginnende co-assistenten. De toets bestaat uit vragen die de belangrijkste gebieden van de Inwendige Geneeskunde bestrijken. Het belangrijkste doel van deze toets is om beginnende co-assistenten, na vaak maanden wachttijd, in de gelegenheid te stellen hun parate kennis te toetsen. Tevens worden zij mogelijk daardoor gestimuleerd zodig hun kennis op te frissen.

De resultaten op deze diagnostische toets vormden de aanleiding om de parate medische kennis bij (beginnende) co-assistenten eens aan een nader onderzoek te onderwerpen en, waar mogelijk, te vergelijken met gegevens van elders. Centrale vraag daarbij was, of er iets te zeggen valt over de vraag of het niveau van deze kennis toereikend mag heten voor het functioneren als co-assistent.

SAMENSTELLING EN SCORING VAN DE ALCO-INTREETOETSEN

Het vragenbestand waaruit de Intreetoetsen worden samengesteld, is opgeslagen op een microcomputer (programmatuur 'TSSA', van het Instituut voor Onderzoek van het Wetenschappelijk Onderwijs, IOWO, Nijmegen). Het bestand bevat momenteel 409 twee- en meerkeuzevragen, gespreid over alle gebieden van de Inwendige Geneeskunde. Hiervan zijn 279 (=68%) vragen afkomstig van vroegere bloktentamens uit de eerste fase van het Nijmeegse geneeskunde-curriculum en 96 (=24%) uit vroegere versies van de Maastrichtse Voortgangstoets. De resterende 34 vragen (=8%) zijn door de ALCO-coördinator speciaal voor deze toets geconstrueerd. Van de beide eerstgenoemde categorieën zijn de vragen geselecteerd door twee internisten, die nauw betrokken zijn geweest bij de ontwikkeling van het ALCO-schap. In het navolgende wordt er dan ook van uit gegaan, dat de Intreetoetsen voldoende validiteit hebben voor een meting van het voor co-assistenten gewenst geachte aanvangsniveau.

Een voorbeeld van het gebruikte type vragen is te vinden in tabel 1.

Telkens als een groep van 12 ALCO-assistenten een Intreetoets moet afleggen (momenteel ongeveer 25 groepen per jaar), wordt uit het vragenbestand een toets van 45 vragen samengesteld.

De toetsen worden gescoord volgens de regel: 1 punt voor een goed antwoord en 0 punten voor een fout of blanco antwoord. Om redenen die te maken hebben met de eigenschappen van dit scoringssysteem, maar waarop in dit bestek niet nader kan worden ingegaan, wordt aan studenten geadviseerd *alle* vragen te beantwoorden. De behaalde score voor de toets is gelijk aan het aantal goede antwoorden. Om een vergelijking met andere gegevens mogelijk te maken werd de behaalde score naar een cijfer op de bekende 10-punts cijferschaal omgerekend. Dit geschiedde via lineaire transformatie, waarbij voor de grens voldoende/onvoldoende een norm van 50% van het reële scorebereik (tussen toevalsscore en totaalscore) en voor het cijfer 10 de maximaal haalbare score wordt aangehouden.

J.C.M. Metz,
coördinator
onderwijsontwikkeling,
Faculteitsbureau Geneeskunde
en Tandheelkunde

J.A. Bulte,
onderwijskundige,
Faculteitsbureau Geneeskunde
en Tandheelkunde

Correspondentieadres:
Faculteitsbureau Geneeskunde
Johannes Wierlaan 3
6500 HB Nijmegen

Tabel 1. Voorbeeld van zes vragen uit de Nijmeegse ALCO-Intreetoets.

1. Het geruis van een mitralisstenose wordt het beste gehoord
 1. over de apex
 2. in de 2e intercostale ruimte rechts van het sternum
2. Een patiënt met een ernstige CARA heeft meer problemen met hoge koorts dan een patiënt met een normale ventilatie
 1. juist
 2. onjuist
3. Indien men bij rectaal toucher melaena vindt, is een bloeding ergens laag in de tractus digestivus het meest waarschijnlijk.
 1. juist
 2. onjuist
4. Bij een acute pancreatitis is er vaak sprake van een défense musculaire.
 1. juist
 2. onjuist
5. Een sterk verhoogd serumgehalte geconjugeerd bilirubine vindt men bij
 1. het syndroom van Gilbert
 2. congenitale hereditaire sferocytose
 3. obstructie-icterus
 4. transfusie met bloed van een verkeerde bloedgroep
6. Tot de predisponerende factoren voor het ontstaan van cholesterolstenen in de galblaas behoort/behoort
 1. het gebruik van orale anticonceptiva
 2. diabetes mellitus
 3. hereditaire sferocytose
 4. 1 + 2 zijn goed

RESULTATEN OP DE ALCO-INTREETOETSEN

Dit artikel heeft betrekking op de resultaten van de ALCO-Intreetoetsen die zijn afgenomen tussen 15 februari 1987 en 1 september 1988. Dit betreft 34 toetsen van elk 45 vragen, die in totaal door 397 studenten zijn gemaakt (34 ALCO-groepen).

De gemiddelde groepsscore (= gemiddeld aantal goed beantwoorde vragen) bedroeg 30.03 (s.d. 1.87). Het laagste groepsgemiddelde bedroeg 26.15 en het hoogste 34.36. Uitgedrukt in procenten: gemiddeld scoorden de groepen 67% van de vragen goed met een maximum gemiddelde van 76% en een minimum gemiddelde van 58%. Volgens de ge-

bruikte transformatieregel komt dit gemiddelde van 67% overeen met het cijfer 5.2 (indien de gemiddelde toevalsscore als beginpunt van de schaal wordt genomen). Van de deelnemende studenten behaalden er 235 (=65%) een onvoldoende. Bij 28 van de 34 ALCO-groepen had meer dan de helft van de studenten uit de groep volgens de bij ALCO gebruikelijke caesuur van 50% van het reële scorebereik, een onvoldoende.

Van meet af aan vonden zowel co-assistenten als tutores deze resultaten nogal teleurstellend. Verondersteld werd, dat deze tegenvallende resultaten mogelijk te maken zouden kunnen hebben met de lange wachttijden tussen het behalen van het doctoraalexamen en het begin van het ALCO-schap. In sommige gevallen waren die namelijk opgelopen tot bijna een jaar. Het zou kunnen zijn, dat de parate kennis, na een aantal weken in het ALCO-schap weer met de leerstof te zijn bezig geweest, als het ware vanzelf weer terugkomt. Om die veronderstelling te toetsen, is door één ALCO-groep ook op de laatste dag van het ALCO-schap een Intreetoets afgelegd. Wat bleek? Alle studenten scoorden toen een onvoldoende! Op de eerste ALCO-dag was dat voor deze groep nog 75%. Er is dus weinig houvast voor de hierboven genoemde verklaringsmogelijkheid.

Was de parate kennis eigenlijk ten tijde van het behalen van het doctoraalexamen wel aanwezig? Ter toetsing van het antwoord op die vraag werden op 7 maart 1988 34 studenten, die op die dag slaagden voor het doctoraalexamen, bereid gevonden een Intreetoets te maken. Van hen scoorden er toen 30 (=80%) een onvoldoende. Uit die bevinding moet geconcludeerd worden, dat kennelijk ten tijde van het behalen van de doctoraalbul veel van de kennis uit eerdere blokken, al is verdwenen.

Er konden derhalve geen aanknopingspunten worden gevonden voor het optreden van kennisvervaging als gevolg van wachttijden, noch voor het wel aanwezig zijn van die kennis aan het begin van de wachttijd. Toch is het aannemelijk, dat die kennis, getuige de herkomst van het merendeel van de vragen, ooit aanwezig is geweest. Alle ALCO-assistenten zijn immers doctorandi, en hebben in eerdere toetsen betere scores op een deel van de vragen behaald.

Nagegaan is nog, of de toetsscores voor de

verschillende onderdelen van de Inwendige Geneeskunde uiteenliepen. Dit bleek niet het geval te zijn. Wel bleken er kleine verschillen te bestaan tussen de scores op vragen van Maastrichtse en van Nijmeegse herkomst de eerstgenoemde bleken de moeilijkste.

VERGELIJKING MET ANDERE GEGEVENS

De vraag dringt zich op hoe deze resultaten afsteken tegen gegevens van elders. In 1984 verscheen in het Nederlands Tijdschrift voor Geneeskunde een studie van Bender et al.¹ naar de medische kennis bij studenten uit de medische faculteiten van Groningen, Leiden en Maastricht. De belangrijkste gegevens uit dat onderzoek zijn samengevat in tabel 2. Bij dat onderzoek werd onder andere gebruik gemaakt van een toets van 64 items, ontleend aan de Maastrichtse Voortgangstoets van mei 1979. Groningse en Leidse studenten maakten deze toets, op basis van vrijwilligheid, in de loop van 1980. Uit de resultaten, die studenten van *alle* studiejaar uit de drie genoemde faculteiten op deze toets hebben behaald, blijkt, dat er onderling nauwelijks verschillen zijn tussen studenten van hetzelfde studiejaar uit de verschillende faculteiten. De kennistoename in de loop van de medische studie geeft in grote lijnen een gelijksoortig beeld te zien bij de drie categorieën studenten.

In de studie van Bender et al. komt de vraag naar het aangetroffen kennisniveau nauwelijks aan de orde. Kunnen de scores die door

de deelnemende studenten op de toets gehaald werden, als toereikend beschouwd worden? Als wij ons beperken tot de 5e en 6e jaars studenten, dan blijkt, dat de gemiddelde toetscores voor de 5e jaars ongeveer 35% bedroegen en voor de 6e jaars 40-45%. De referentiegroep van Maastrichts huisartsen-in-opleiding scoorde gemiddeld iets hoger dan 50%. Bij de interpretatie van deze gegevens dient de lezer zich te realiseren, dat de betreffende toetsen gescoord zijn volgens een regel die afwijkt van de regel, die boven voor het Nijmeegse ALCO-schap is beschreven. De door Bender et al. gebruikte steekproef uit de Maastrichtse Voortgangstoets wordt gescoord volgens een regel waarbij negatieve punten worden gegeven voor onjuiste antwoorden. De toetsscores worden berekend door van het aantal goede antwoorden de foutscore af te trekken. Bij omrekening van deze scores naar een cijfer van de 10-puntsschaal, dient men zich te realiseren, dat het scorebereik bij dit systeem afwijkt van dat bij het eerder beschreven systeem, waarbij immers de toevalsscore het begin van het scorebereik vormt. Wordt bij de gegevens van Bender et al. een analoog criterium van 50% voor de grens voldoende/onvoldoende aangehouden als boven beschreven voor de ALCO-toetsen, dan moet de conclusie luiden, zoals blijkt uit tabel 2, dat de gemiddelde prestaties van de 5e en 6e jaars studenten van alle drie de faculteiten beneden dat criterium blijven.

In de loop van 1983-1984 is met een representatieve groep Nijmeegse studenten uit alle

Tabel 2. Samenvattend overzicht van de in dit artikel gerefereerde onderzoeken, waarbij de toetsscores lineair zijn getransformeerd naar cijfers. De grens voldoende/onvoldoende is bij systeem 1 gelegd bij 50% van het reële scorebereik en bij systeem 2 bij 50% van de totaalscore.

(d)MVT = (deel van) Maastrichtse Voortgangstoets.

plaats/tijd	ALCO KUN '88		MVT KUN '83		MVT RULi '83		dMVT RULi '79		dMVT RUL '80		dMVT RUG '80	
studiejaar	5	6	5	6	5	6	5	6	5	6	5	6
gem.score (systeem 1)	67	-	-	-	-	-	-	-	-	-	-	-
gem.score (systeem 2)	-	-	31	40	34	42	35	-	35	42	34	44
cijfer	5.2	-	3.6	4.6	3.9	4.8	4.0	-	4.0	4.8	3.9	5.0

studiejaren een soortgelijke meting als beschreven door Bender et al. gedaan. Tabel 2 bevat de gegevens van de groep 5e (N= 27) en 6e (N= 21) jaars, die de gehele Maastrichtse Voortgangstoets van maart 1983 (264 vragen) maakten.

Hun resultaten werden vergeleken met die van hun Maastrichtse jaargenoten (resp. 65 en 54 co-assistenten). Als referentiegroep fungeerde een groep van 42 arts-assistenten, die een gemiddelde score van 44% behaalde. De Nijmeegse 5e jaars scoorden gemiddeld 31% en de 6e jaars 40%. Voor Maastricht waren deze percentages resp. 34 en 42. Ook hier luidt de conclusie, dat deze prestaties over de gehele lijn van dezelfde orde zijn en beneden het eerder aangehouden criterium blijven.

Uit een recent gepubliceerde studie van Bulte et al.² naar de medische kennis bij huisartsen-in-opleiding, komt naar voren, dat de grootteorde van de gemiddelde scores op de Maastrichtse Voortgangstoets (en op de subtoets Huisartsen daaruit) bij huisartsen-in-opleiding niet van het boven aangeduide niveau afwijkt, ongeacht de Nijmeegse dan wel Maastrichtse herkomst van de arts-assistenten. Overigens blijkt de kennis, zoals gemeten met deze toets, tijdens de huisartsopleiding wel significant te stijgen.

BESCHOUWING

De beschreven resultaten van co-assistenten en arts-assistenten op kennistoetsen op het niveau van resp. beginnende co-assistenten en afgestudeerde basisartsen, suggereren, dat het met die kennis maar mager gesteld is. In ieder geval moet men vaststellen, dat er kennelijk op dit type toetsen nauwelijks hogere scores te behalen zijn. Een aantal kanttekeningen moet echter worden gemaakt.

In alle gerefereerde gevallen hebben de deelnemers de toetsen onvoorbereid afgelegd. Dit kan een neerwaarts effect op de scores hebben gehad, maar daardoor ontstaat wel een zuiverder beeld van het parate kennisniveau. Indien dit effect is opgetreden, dan vermoedelijk voor allen in gelijke mate. Overigens, op de Voortgangstoets is een gerichte voorbereiding nauwelijks mogelijk. Deze vindt ook in de normale toepassingssituaties niet plaats. Met uitzondering van de Maastrichtse studenten ging het voor alle deelnemers om 'niet-

meetellende' toetsen: hetzij omdat deelname vrijwillig was, hetzij (zoals in het geval van de ALCO-Intreetoets) omdat de toets alleen diagnostische betekenis heeft. Moeilijk te beantwoorden is de vraag of, en hoe, dit van invloed is geweest op de toetsresultaten. Wij neigen tot de veronderstelling, dat deelname aan niet-meetellende toetsen in ieder geval niet tot een verhoging van de scores leidt. Indien men het gedrag van studenten die aan dergelijke toetsen deelnemen, observeert, is er weinig reden om aan te nemen dat dit niet serieus gebeurt. In vele gevallen is de keuze tot meedoen vrijwillig, waardoor een zekere selectie optreedt: de deelnemers zijn echt geïnteresseerd in wat zij feitelijk zullen blijken te weten.

Bij de interpretatie van de resultaten moet ook rekening worden gehouden met de mogelijke effecten van verschillende scoringsystemen. De ALCO-Intreetoetsen worden, zoals boven uiteengezet, op een andere wijze gescoord dan de Maastrichtse Voortgangstoets. Deze scoringsystemen veronderstellen een onderling verschillende invulstrategie bij vragen waarop men het antwoord niet weet. De Maastrichtse deelnemers mogen vertrouwd geacht worden met het bij de voortgangstoetsen gebruikelijke scoringsstelsel, de Groningse en Leidse studenten niet. De Nijmeegse ALCO-Intreetoets wordt gescoord volgens een systeem, waarmee Nijmeegse co-assistenten niet vertrouwd zijn, omdat in de eerste fase van hun studie met het andere, ook in Maastricht gebruikelijke systeem gewerkt wordt. De Nijmeegse ALCO-studenten worden bij de Intreetoets over de optimale invulstrategie geïnformeerd. Berekent men voor de Groningse en Leidse studenten het percentage goed beantwoorde vragen (en niet goed minus foutscore) als maat voor hun score, dan komt men voor de verschillende categorieën deelnemers tot percentages tussen de 50 en de 70. Ook deze cijfers liggen beneden het eerder aangehouden criterium.

Uiteen analyse van de evaluatiegegevens van het Nijmeegse ALCO-schap blijkt, dat studenten aan het einde van hun ALCO-schap terugkijkend op hun aanvangsniveau, van oordeel zijn, dat hun kennis ontoereikend was. Dit vormt een argument voor de validiteit van de resultaten op de Intreetoets. Zijn er verklarende factoren voor de beschreven bevindingen, die kennelijk niet tot één

faculteit beperkt zijn? Deugen de gebruikte toetsen niet? Dit lijkt onwaarschijnlijk, gezien de zorgvuldigheid waarmee deze toetsen worden samengesteld. Daarenboven is met betrekking tot de Maastrichtse toets zowel door een groep Maastrichtse als door een groep Nijmeegse artsen 75% van de vragen als relevant gekwalificeerd.

Zijn co-assistenten de in de eerste vier jaar van hun studie aangeleerde kennis misschien weer vergeten? Leren zij de voor hun functioneren als arts noodzakelijke onderwerpen verkeerd? Of leren zij de verkeerde onderwerpen? Over het antwoord kan slechts gespeculeerd worden.

In de beschreven onderzoeken ging het steeds om kennis. Medische (feiten)kennis is weliswaar een noodzakelijke, maar niet voldoende voorwaarde voor medische competentie. Helaas zijn er geen gegevens beschikbaar van onderzoeken, waarin studenten van verschillende faculteiten op verschillende factoren van medische competentie worden vergeleken. Daaruit kan de aanbeveling voortvloeien op grotere schaal een vergelijkend longitudinaal onderzoek op te zetten met behulp van voortgangstoetsen - ook op andere dan kennisdomeinen. Dergelijke toetsen zijn, beter dan de gebruikelijke vak- of bloktoetsen, in staat de ontwikkeling in kennis en vaardigheden te registreren en te stimuleren.

LITERATUUR

1. Bender W, Cohen-Schotanus J, Imbos T, Versfelt WA, Verwijnen AM. 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, Van Leeuwen YD. Kennis bij huisartsen in opleiding; kennismeting bij Nijmeegse arts-assistenten. Medisch Contact 1988; 43: 1426-8.

In wetenschappelijke kringen zijn begrippen als "citation index" en "impact factor" ingeburgerd. Rondom deze begrippen hangt een wetenschappelijk aureool. Mijs inziens zijn er ook andere begrippen wenselijk om publicaties te karakteriseren. Van een "leesbaarheidsfactor" heb ik bijvoorbeeld nog nooit gehoord. Ook mis ik een "relevantie-index", waarin de relevantie voor bijvoorbeeld het medisch handelen of het medisch onderwijs wordt uitgedrukt. Sommigen zien wellicht een verband tussen de "relevantie-index" en acceptatie van een artikel door een tijdschrift met een hoge "impact factor". Voor zover mij bekend is dat verband echter niet wetenschappelijk aangetoond. Een onderzoek naar leereffecten van in wetenschappelijke kringen hoog aangeschreven publicaties lijkt mij ook niet onbelangrijk. Kortom, ik twijfel aan de validiteit en de betrouwbaarheid van het etiketteren van publicaties zoals dat nu gebeurt. Als dat etiketteren ook nog consequenties heeft voor mensen, beschouw ik het als een in vitro experiment dat ten onrechte in vivo wordt uitgevoerd. Het oordeel van een facultaire medisch ethische commissie over een dergelijk experiment zou 'onvoorwaardelijk negatief' luiden.

A.J.J.A. Scherpbier