

DE BEKWAAMHEID VAN DE ARTS:

WAT WETEN WE ERVAN EN WAT KUNNEN WE ER VAN WETEN?

Dr. H.J.M. van Rossum
Hoofd afdeling Onderwijs en
Onderzoek
Faculteit der Geneeskunde
Rijksuniversiteit Groningen

Correspondentieadres:
Bloemsingel 1
9713 BZ Groningen

Op het moment, dat een co-assistent haar handtekening zet onder de bul, is zij bevoegd de genees-, heel- en verloskunde in Nederland, en in principe binnen de EG, in volle omvang uit te oefenen. Iedere burger mag verwachten dat, voorafgaande aan de verlening van een zo ingrijpende bevoegdheid, een gedegen onderzoek heeft plaatsgehad naar de bekwaamheid van de arts. De toegankelijke neerslag van dit onderzoek vormt de cijferlijst, zoals die tijdens de zitting van de artsexamencommissie ter besluitvorming voorligt. Beziert men een dergelijke cijferlijst, dan kan men zich afvragen wat de betekenis is der cijfers. Voor een patiënt is vooral de prognostische waarde van vitaal belang. Wat betekent bijvoorbeeld een zes voor verloskunde? Is dit de gemiddelde Apgarscore van de eerste tien bevallingen die de jonge arts zal gaan begeleiden? Of weet de jonge arts net voldoende van alle problemen om geen brokken te maken? Of kan zij 60 procent der problemen adequaat oplossen en 40 procent niet? De 'zes' maakt duidelijk, dat zij niet volmaakt is, anders stond er een 'tien'. Hoe gaat de arts om met problemen die zij blijkbaar niet of niet volledig beheerst? Deze vragen zullen zeker opkomen in het hoofd van een patiënt, indien in de wachtkamer niet het diploma, maar de cijferlijst van de pas afgestudeerde arts hangt.

RELATIE TUSSEN DIPLOMA EN CIJFERLIJST

Wat is de betekenis van een cijfer? Laten we eens kijken hoe bijvoorbeeld bovengenoemde 'zes' totstandkwam.

De kandidaat onderzocht een patiënt die drie weken tevoren was opgenomen vanwege zwangerschap bij habituele abortus. De co-assistent beschreef alle bevindingen en analyses in de status. Deze status werd vervolgens besproken met de examiner. Dit alles verliep vlot en zonder problemen. Na deze bespreking stelde de examiner nog enkele vragen over andere ziektebeelden, waaronder diabetes in de zwangerschap. Na een antwoord van de kandidaat over de te verrichten diagnostiek, ontplofte de examiner met de mededeling, dat de kandidaat beslist geen voldoende kon krijgen vanwege dat ene antwoord. ("Ik geef in beginsel nooit hoger dan

een acht, maar deze fout is niet toelaatbaar"). Het protest, dat de co-assistent na afloop indiende bij de vakgroepvoorzitter, leidde tot het inwinnen van nadere informatie bij de zaalassistenten die de 'co op zaal hadden meegemaakt'. Hun verklaring, dat de co-assistent een prettige collega was geweest die goed de patiënten op hun gemak kon stellen, leidde ertoe dat een zes als eindcijfer werd vastgesteld. De co-assistent liet het erbij, omdat zij geen aspiraties had voor een 'cum laude' en omdat ze al een toezegging had snel met de opleiding tot oogarts te kunnen beginnen.

Dit verhaal is uiteraard niet maatgevend voor alle beoordelingen in Nederland, maar had wel gisteren ergens in den lande kunnen gebeuren. Met deze illustratie wil ik aantonen, dat er een discrepantie bestaat tussen de omvang en implicaties van de bevoegdheid en de omvang en zorgvuldigheid van het onderzoek, dat tot het verlenen van die bevoegdheid leidt. Tussen diploma en cijferlijst gaapt een kloof. De kwaliteit van het artsexamen is onder de maat.¹

AANVANGSSITUATIE BEROEPSOPLEIDINGEN

Tot zo'n twintig jaar geleden waren de werkzaamheden van de huisarts maatgevend voor de inrichting van het artsexamen. Daarna is het anders, maar niet zoveel moeilijker geworden om een indicatie te krijgen over de eisen die aan de basisarts gesteld moeten worden. De meeste als arts werkzame basisartsen zullen immers in het begin van hun verdere carrière tenminste een deel van hun tijd besteden aan direct contact met patiënten (anamnese, lichamelijk onderzoek, diagnostiek, communicatie, uitleg, ingrepen), een deel aan schriftelijke verslaglegging en correspondentie (status, brieven, formulieren), een deel aan mondelinge intercollegiale communicatie (presentaties, overdracht, consulten) en een deel aan eigen studie (deels gekoppeld aan eigen patiënten, deels als voorbereiding van referaten of nascholingsbijeenkomsten). Wanneer men op grond van de resultaten van het artsexamen goed wil voorspellen hoe de kandidaat in de volgende fase van zijn opleiding zal functioneren, dan ligt het voor de hand het artsexamen in te richten op geleide van die opleiding.

Het bovengenoemde scala van activiteiten kan worden gegroepeerd en samengevat in verschillende categorieën, zoals de categorie klinische vaardigheden, de categorie informatie-overdracht en de categorie eigen functioneren.

AFLEIDEN INHOUD ARTSEXAMEN UIT EEN TOETSMATRIJS

De genoemde indeling in categorieën is enigszins kunstmatig en schematisch, maar kan worden gebruikt als mogelijk uitgangspunt bij het ontwerpen van een toetsmatrijs voor een systematische inrichting van het artsexamen.

Een ander uitgangspunt kan worden ontleend aan de empirische probleemoplossingscyclus.² Ter toelichting daarop het volgende. In bepaald opzicht kan het artsexamen gekarakteriseerd worden als een situatie waarin professioneel menselijk handelen moet worden beoordeeld. Dit handelen kan aangeduid worden als het handelen in probleemsituaties. De empirische cyclus bestaat erin, dat een mens, geplaatst in een probleemsituatie (P) zich eerst een oordeel vormt over dat probleem (O) en vervolgens een programma van activiteiten (P) ontplooit, dat de probleemsituatie beïnvloedt. Deze cyclus blijft motor van handelen tot het probleem is opgelost. De drie onderdelen (POP) van deze empirische cyclus kunnen als een tweede uitgangspunt worden gebruikt bij het ontwerpen van een toetsmatrijs voor het artsexamen.

Door de onderdelen van de empirische cyclus te kruisen met de eerder genoemde activiteiten-categorieën van de arts ontstaat een toetsmatrijs (tabel 1) met negen elementen, waaruit per examen op systematische wijze een representatieve steekproef kan worden getrokken.

Tabel 1.
Toetsmatrijs

I. PROBLEEMSITUATIE	Voorbeelden
1. kliniek	patiënt
2. informatie-overdracht	gezondheidszorg-systeem
3. eigen functioneren	arts zelf
II. OORDEELSVORMING	
1. kliniek	diagnose, prognose
2. informatie-overdracht	bias bij consult, medicalisering
3. eigen functioneren	intercollegiale toetsing
III. PROGRAMMA	
1. kliniek	anamnese, onderzoek, ingrepen, troost
2. informatie-overdracht	consult vragen, verwijzen, wetten
3. eigen functioneren	(bij)leeractiviteiten, consensus

Voor de noodzakelijke inhoudelijke specificering van de eisen voor het artsexamen is een inventarisatie nodig van clusters van relevante ziektebeelden, symptomen en andere onderwerpen, die daarvoor als leidraad kan dienen. Behalve de inhoud van het artsexamen is ook de vorm van belang. Voor ieder inhoudelijk cluster zal een beoordelingssituatie moeten worden ontworpen. Voor technische zaken als schaalkeuze en criteriumvaststelling is samenwerking tussen de professie en onderwijskundigen onontbeerlijk. Hieronder volgt een kritische bespreking van het huidige artsexamen aan de hand van bovenstaand schema.

PROBLEMSITUATIES

Op het eerste gezicht lijkt het huidige artsexamen adequaat. Een groot deel ervan bestaat uit beoordelingen van stages; echte praktijk dus. De co-assistent werkt met patiënten en haar werkzaamheden worden beoordeeld door leermeesters, die in de directe omgeving aanwezig zijn. Alle drie de elementen - patiënt, gezondheidszorgsysteem en arts - zijn vertegenwoordigd.

Dat adequaat zijn is echter maar schijn, want het probleem is, dat de werkzaamheden en de rol van de co-assistent meestal slecht gedefinieerd zijn, dat het beoordelingsproces niet of lacunair geprotocolleerd is en dat de leermeesters niet getraind en vaak niet geschikt zijn om te beoordelen.

In de stage zijn ook enkele meer gestructureerde beoordelingsmomenten opgenomen in de vorm van patiënt-gebonden examens bij een examiner, vaak van buiten de afdeling. De inrichting van deze examens staat echter veelal ver van de praktijksituatie. Er vindt een gesprek plaats naar aanleiding van een status of patiëntenprobleem; het betreft dus meer spreken over de praktijk dan beoordelen in de praktijk.

Veel praktijksituaties zijn in het huidige artsexamen meestal niet vertegenwoordigd, zo-

als het aanvragen van een consult, het verwijzen, het geven van uitleg aan de patiënt, het voeren van een troostgesprek, het inlichten van de familie en het voeren van overleg met andere werkers in de gezondheidszorg, zoals verpleegkundigen. Blijkbaar wordt er van uitgegaan, dat deze vaardigheden in de vorm van genade worden geschonken tijdens het signeren van de bul, want verdere training en beoordeling vinden niet plaats en na het artsexamen moet de arts deze dingen wel doen. Ten aanzien van de inhoud van het examen zou men verwachten, dat de problemen die in het examen aan de orde komen, een spreiding vertonen over parameters als ziektebeelden, geslacht, leeftijd, acuut of chronisch beloop en mate waarin het leven wordt bedreigd. De inhoud wordt echter niet bewaakt.

Het beoordelen van het eigen functioneren kent enkele bijzondere aspecten, met name de perceptie door de persoon zelf en de rol die de omgevingscultuur speelt bij het beoordelen. Het is niet eenvoudig het persoonlijk functioneren te beoordelen. Meestal gebeurt het dan ook niet, hoewel het belang van 'een goede attitude' ten aanzien van het eigen functioneren vaak wordt beleden. Nodig is, dat in de werksituatie op de afdelingen een georganiseerde vorm van intercollegiale toetsing bestaat, die is geaccepteerd als onderdeel van kwaliteitsbevorderende maatregelen.

OORDEELSVORMING

Deze categorie onttrekt zich aan directe beoordeling, waardoor men snel in de mystieke sfeer komt van 'klinische blik', 'dit leer je in de praktijk' en dergelijke. Met de toenemende aandacht van ethici en vooral psychologen, wier terrein dit is, komt er echter enig zicht op een beoordelingsinstrumentarium. Een eerste begin is gemaakt, hoewel veel nog ontbreekt. Van oudsher wordt veel aandacht besteed aan de beoordeling van het diagnostisch oordeel in een patiëntencasus. De kern van de beoordeling vormt meestal echter de motivering achteraf van de differentiële diagnose. Onderzoek van de laatste jaren richt zich meer op de inhoud van het oordeel zelf en op de wijze waarop dit oordeel tot stand komt. Dit oordeel komt tot stand door associaties en wordt beïnvloed door eerdere ervaringen met patiënten.³

Er worden momenteel meetinstrumenten ontwikkeld, die geschikt lijken het differentieel-diagnostisch oordeel zelf vast te leggen.⁴

⁵ Gepoogd wordt op die manier meer tot de kern te komen, omdat het principe van die meetinstrumenten bestaat in een beargumentatie van de in het werkgeheugen terecht gekomen diagnostische hypothesen. Dergelijke meetprocedures kosten weinig tijd, waardoor meerdere diagnostische terreinen in de meting kunnen worden opgenomen.

Het beoordelen van procedures en werkwijzen in het zorgsysteem en het beoordelen van het eigen functioneren zijn niet in het arts-examen opgenomen. Beide zijn niettemin van groot praktisch belang en komen in toenemende mate in de belangstelling. De onderwerpen zijn wat heterogeen en omvatten naast onderwerpen als asepsis en antibiotica-beleid ook de waarde van thuiszorg en de appreciatie van teamwork, de invloed van eigen normen op het functioneren en het geven en verwerken van kritiek op het functioneren.

Ik noem enkele voorbeelden van onderwerpen die recent meer in de belangstelling zijn gekomen en doorgaans nog niet in het arts-examen zijn doorgedrongen. De wijze waarop een consult wordt aangevraagd voor het beoordelen van ECG's of Röntgenfoto's kan een foutenbron creëren, waardoor het consult mogelijk aan waarde inboet. De wijze waarop laboratorium-uitslagen worden gepresenteerd is van invloed op het memoriseren ervan door artsen.⁶ Het daadwerkelijk handelen wordt mede bepaald door de inrichting van de werksituatie, die soms in conflict is met regels en protocollen die voor het werk zijn ontworpen.² Een duidelijk voorbeeld is de statusvoering. Co-assistenten worden getraind aan de hand van een modelstatus, die gebaseerd is op de probleemgeïntegreerde werkwijze. Eenmaal op zaal maakt de heersende praktijk het hun moeilijk deze werkwijze verder te ontwikkelen. Eigen normen en waarden hebben invloed op de beoordeling van de toestand van een patiënt. In de onderwijssituatie komt dit sterk naar voren tijdens het onderdeel 'Ethiek' in het Leidse Alcoschap. Hierin wordt in een socratische dialoog de invloed van eigen normen onderzocht op het beoordelen van hypothetische patiënten.

Als conclusie mag gelden: er is werk aan de winkel voor personen die het geweten van het zorgsysteem (moeten) vormen.

PROGRAMMA

Het formuleren van een programma waarin klinische vaardigheden zijn opgenomen, is als onderdeel goed vertegenwoordigd in het huidige artsexamen. Het betreft echter het formuleren van wat men zou gaan doen. De uitvoering zelf wordt meestal niet geobserveerd, hoewel het instrumentarium hiervoor aanwezig is. In de cursussen systematische vaardigheidstraining zijn goede trainings- en meetprocedures ontwikkeld aan het begin van de co-assistentenschappen.^{7,8} Het resultaat aan het einde van de co-assistentenschappen blijft vaak onbeoordeeld. Tijdens de co-assistentenschappen vindt zelden beoordeling plaats door middel van gestructureerde directe observatie.

Er is dus een 'dekkingsprobleem'. Het instrumentarium is er wel, maar het ontbreekt aan tijd en organisatie - dus aan geld en goede wil - om de noodzakelijke beoordelingen uit te voeren. Dat het noodzakelijk is, blijkt uit de discrepantie tussen de eisen van docenten, zoals die in het BOC-rapport zijn neergelegd, en de resultaten van onderzoek, zoals verricht door Tan.¹⁹ Communicatie met werkers in de gezondheidszorg neemt een prominente plaats in doordat de statussen en brieven een belangrijke component zijn van de huidige beoordelingsprocedures. Het is echter vaak de inhoud die wordt beoordeeld, meer dan de vorm, waar het hier om gaat. Patiëntenpresentaties vormen wel onderdeel van training gedurende de co-assistentenschappen, maar het beoordelen ervan gebeurt niet systematisch en er is geen registratie van de beoordeling. Communicatie met verpleegkundigen wordt meestal niet betrokken bij het beoordelen van de co-assistent.

Het vermogen om deficiënties in kennis en vaardigheden aan te vullen is niet opgenomen in het huidige artsexamen, het vermogen om kritiek van anderen (peer review, functioneringsgesprekken, patiëntenkritiek) om te zetten in leer-activiteiten evenmin.

WAARDE VAN EXAMENRESULTATEN

Resultaten van metingen in een kunstmatige situatie mogen niet zondermeer worden generaliseerd naar de werkelijke situatie. Iemand's 'competence' (gedrag in een meet-situatie) is niet gelijk aan iemand's 'performance' (gedrag in de praktijksituatie, waarin men zich onbeoordeeld waant). De validiteit van de meting is in het geding.

Een opmerkelijke studie is onlangs verricht door Rethans et al.¹⁰ Hij trainde proefpersonen om een klacht te spelen. Vervolgens werden deze personen in het gewone spreekuur van een aantal huisartsen binnengesmokkeld, zonder dat deze wisten welke het waren. De 'performance' van de huisartsen werd gescoord door de daartoe getrainde patiënten. Enkele maanden later werden dezelfde artsen uitgenodigd voor een spreekuur-examen, waarin zij dezelfde klachten kregen voorgeschoteld gespeeld door andere patiënten. De 'competence' van de artsen bij die gelegenheid werd op dezelfde wijze gescoord als de eerste keer. Na afloop werden beide scores vergeleken. Voor alle problemen gold, dat de scores in de tweede sessie hoger waren. Van meer belang echter was de bevinding, dat de correlatie tussen beide scores nul was. Daaruit kan de conclusie worden getrokken, dat gedrag in een testsituatie niet het gedrag in de eigen praktijk voorspelt.

SAMENVATTEND

Het huidige artsexamen bestaat voornamelijk uit het beoordelen van kennis die gerelateerd is aan patiëntenproblemen, in een weinig tot niet gestructureerde setting van gesprekken met examinatoren. Het artsexamen kent ten minste drie problemen:

1. Een dekkingsprobleem

Hoewel er goed meetinstrumentarium aanwezig is, komen met name factoren als bekendheid met ziektebeelden en het beheersingsniveau van medische vaardigheden niet representatief aan de orde.

2. Een instrumenteel probleem

Onderdelen die zeer belangrijk gevonden worden (diagnostisch oordeel, zelfbeoordeling, omgaan met deficiënties) kunnen niet goed worden beoordeeld, omdat er geen goede technieken voor zijn. Ondanks vorderingen op het terrein van het meten van diagnostische oordeelsvorming, is verder onderzoek en ontwikkelingshulp van psychologen nodig om in deze leemte te voorzien.

3. Een validiteitsprobleem

Aangetoond is, dat metingen in een examensituatie niet correleren met metingen in de praktijk. Verder onderzoek is nodig om voor dit probleem een oplossing te vinden. Het zou wel eens kunnen zijn, dat een aantal omissies in de beoordeling bij het huidige artsexamen pas ongedaan kan worden gemaakt als het fenomeen 'beoordeling' (peer review, functioneringsgesprekken, visitatie) een wezenlijk en geïntegreerd deel van de medische praktijkvoering is gaan uitmaken. Een voorbeeld-gevende en initiërende rol is in dit opzicht weggelegd voor het afdelingshoofd, de opleider der co-assistenten.