

Richtlijn interpretatie en gebruik van toetsresultaten in het portfolio

M.J.B. Govaerts, E. Driessen, B. Verhoeven, C.P.M. van der Vleuten, H. Brackel, J. van Hoorn, R. van de Laar, J. Maas, S.G. Oei

Voorwoord

Deze Richtlijn in het kader van het In VIVO-project is geschreven door drs. Marjan Govaerts, namens het kernteam In VIVO van de OOR Zuid-Oost Nederland (OOR-ZON), met medewerking van dr. Erik Driessen, dr. Bas Verhoeven, prof. dr. Cees van der Vleuten (vakgroep Onderwijsontwikkeling en Onderwijsresearch, Faculty of Health, Medicine and Life Sciences, Universiteit Maastricht) en van dr. Hein Brackel (voorzitter), drs. Jeroen van Hoorn, drs. Rafli van de Laar, dr. Jacques Maas, en prof. dr. Guid Oei (kernteam In VIVO OOR-ZON, werkgroep Toetsing).

De overige leden van het kernteam In VIVO OOR-ZON, dr. Jamiu Busari (voorzitter), prof. dr. Gerard Essed, drs. Florian ten Cate, dr. Twan Mulder, drs. Mark Otting en dr. Frans Roumen, zijn hierbij geconsulteerd.

Status van de Richtlijn

De Richtlijn is gebaseerd op literatuuronderzoek, best practices en ervaringen van de auteurs en heeft een adviserend karakter.

Deze Richtlijn dient te zijner tijd te worden herzien op basis van aanvullend onderzoek, ervaringen van gebruikers van de Richtlijn en op basis van bredere ervaring met de implementatie van de herziene opleidingsplannen medisch-specialistische vervolgopleidingen. Naar verwachting zal een eerste herziening-update van de Richtlijn twee jaar na grootschalige implementatie plaatsvinden.

Inleiding

Gedurende de opleiding maken artsen in opleiding tot medisch specialist (aios) en hun opleiders gebruik van een portfolio om het leerproces van de aios te volgen, waar nodig bij te sturen en te beoordelen. Toetsresultaten vormen een essentieel onderdeel van de informatie in dit portfolio. De Richtlijn "Interpretatie en gebruik van toetsresultaten in het portfolio" geeft een samenvattend overzicht van aandachtspunten bij interpretatie en gebruik van toetsresultaten in de opleiding tot medisch specialist.

Doel van de Richtlijn

De Richtlijn is een hulpmiddel voor aios en opleiders bij de interpretatie en gebruik van toetsresultaten gedurende de opleiding tot medisch specialist en bevat een aantal praktische handvatten en stappenplannen voor het gebruik van de toetsresultaten in begeleidingsgesprekken (voortgangs- en beoordelingsgesprekken). Hierbij wordt niet alleen aandacht besteed aan de functie(s) die toetsresultaten kunnen vervullen bij het opleiden van aios maar ook aan kwaliteitseisen die aan toetsresultaten en toetsystemen gesteld moeten worden ten behoeve van geloofwaardige en verdedigbare besluitvorming. De Richtlijn bevat daarmee informatie die ook voor andere doelgroepen relevant kan zijn: leden van opleidingsteams, opleidingsmanagement, beleidsmakers en medewerkers kwaliteitszorg.

Voor aios en opleider (-teams) staan in de Richtlijn de belangrijkste aandachtspunten

bij het gebruik van toetsresultaten (werkplekbeoordelingen) in voortgangs- en beoordelingsgesprekken beschreven. Aan de hand van de geschetste criteria kan een inschatting gemaakt worden van de kwaliteit en bruikbaarheid van toetsresultaten, en kan de zorgvuldigheid in besluitvorming getoetst worden.

Voor *opleiders en opleiderteams* worden daarnaast een aantal richtlijnen gegeven voor de inrichting van de toetsing op de werkplek en voor de procedures ten behoeve van besluitvorming.

Voor *opleidingsmanagers en kwaliteitszorg-medewerkers*, tenslotte, staan in de Richtlijn de belangrijkste randvoorwaarden voor een goed toetsstelsel. De Richtlijn is daarmee te gebruiken als basis voor een interne visitatie naar de kwaliteit van het toetsstelsel op een afdeling c.q. binnen een instelling.

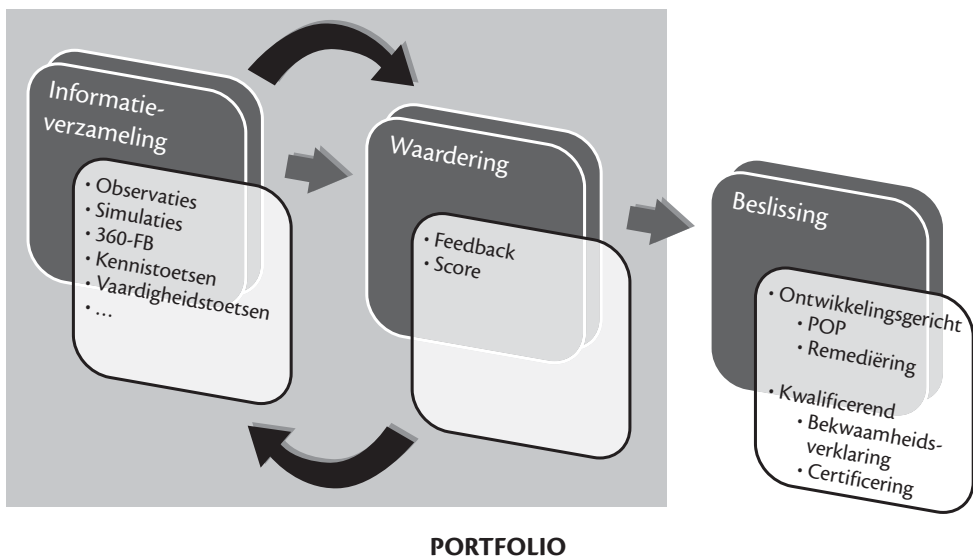
Leeswijzer

De Richtlijn “Interpretatie en gebruik van toetsresultaten in het portfolio” bestaat uit

twee onderdelen. Het eerste deel van de Richtlijn is bedoeld voor toepassing in de dagelijkse opleidingspraktijk. Een aantal belangrijke uitgangspunten en aandachtspunten bij toetsing wordt hierin samengevat. Het tweede deel bevat de uitwerking en onderbouwing van de Richtlijn, op basis van theoretische inzichten en praktijkervaringen in verschillende contexten, waaronder het medisch opleidingscontinuüm. Dit deel wordt afgesloten met een aantal kaders en stappenplannen die een overzicht geven van de aandachtspunten en randvoorwaarden bij toetsing in de medisch-specialistische vervolgopleiding.

Toetsing in de medisch-specialistische vervolgopleiding

Toetsing tijdens de medisch-specialistische vervolgopleiding kan gedefinieerd worden als ‘systematische en structurele evaluatie van het functioneren van de aios voor formatieve en/of summatieve doeleinden’ (zie Figuur 1). De belangrijkste functie van toetsen is het sturen van het



Figuur 1. Het proces van toetsing zoals zichtbaar gemaakt in het portfolio.

leren en het stimuleren van de competentieontwikkeling van de aios, door het verschaffen van feedback, en het vaststellen van sterke kanten en verbeterpunten in het functioneren (formatieve functie van toetsen; ontwikkelingsgericht). Deze formatieve functie van toetsing is belangrijk bij alle vormen van toetsing in de vervolgopleiding, maar vooral bij werkplekbeoordelingen staat de ontwikkelingsgerichte functie centraal. Concrete, specifieke en frequente feedback op het functioneren in de praktijk stelt de aios in staat gericht actie te ondernemen om de kwaliteit van handelen te verbeteren. Documentatie van de feedback in het portfolio – in de vorm van bijvoorbeeld Korte Praktijkbeoordelingen (KPB), Objective Structured Assessment of Technical Skills (OSATS) of 360 graden feedback – stelt aios en opleiders in staat het leerproces te monitoren en te sturen.

Toetsresultaten in het portfolio vormen daarnaast ook bewijsmateriaal dat gebruikt kan worden bij het onderbouwen van beslissingen over bekwaamheid en voortgang in de opleiding (summatieve functie van toetsen; kwalificerend). Documentatie van feedback en evaluatie van performance is dan essentieel bij het tijdig signaleren van aios die in aanmerking dienen te komen voor extra begeleiding of remediëring.

Toetsing tijdens de medisch-specialistische vervolgopleidingen is dus gericht op het zichtbaar maken van het functioneren van de aios, gerelateerd aan de doelstellingen van de opleiding. De doelstellingen voor de opleiding als geheel en voor de verschillende opleidingsfasen zijn beschreven in het opleidingsplan, opgesteld door de betreffende wetenschappelijke vereniging, en uitgewerkt in het regionaal opleidingsplan. De opleidingsdoelstellingen hebben zowel betrekking op algemene competenties (CanMEDS) als op

specialisme-gebonden competenties (of bekwaamheden, geordend in thema's).

Toetsen van competenties

In het algemeen geldt dat betrouwbare en valide uitspraken over competenties alleen mogelijk zijn op basis van gevarieerde informatie, afkomstig van verschillende informatiebronnen (verschillende toetsinstrumenten, verschillende beoordelaars). Er zijn binnen de medisch-specialistische vervolgopleiding vele toetsmethoden en -instrumenten beschikbaar (zie Tabel 1). Elk van de instrumenten levert specifieke informatie over deelaspecten van het professionele handelen van de aios, vaak gekoppeld aan meerdere competentiedomeinen.

(Frequente) beoordelingen van het functioneren op de werkplek (het 'DOES-niveau' van Miller) en expertoordelen van de leden van het opleiderteam zijn essentieel voor het beoordelen van competenties in de medisch-specialistische vervolgopleiding. Relevante informatie over het functioneren van een aios in een bepaald competentiedomein wordt echter ook verschaft door kennistoetsen, simulatietoetsen, wetenschappelijke activiteiten, gevolgde scholing en cursuscertificaten, en registratie van praktijkervaringen (logboek).

Richtlijnen en minimumeisen voor toetsing van professionele competenties zijn beschreven in de (regionale) opleidingsplannen. Uit onderzoek blijkt dat voor het nemen van betrouwbare beslissingen *veel toetstijd* noodzakelijk is. Voor kennistoetsen (KNOWS HOW) geldt bijvoorbeeld dat 2-3 uur toetstijd vereist is voor het kunnen doen van betrouwbare uitspraken; voor het SHOWS HOW-niveau is gemiddeld 3-4 uur toetstijd noodzakelijk (op jaarbasis, eventueel verspreid in de tijd af te nemen). Voor het DOES-niveau gelden vergelijkbare richtlijnen. Concreet betekent dit dat een aios voor het DOES-niveau op jaarbasis bijvoorbeeld acht KPB's,

Tabel 1. Overzicht van toetsmethoden en -instrumenten naar Miller-niveau.

Miller-niveau	Toetsmethoden / -instrumenten
DOES	KPB, 360 graden feedback; OSATS, opdrachtverslagen (bijvoorbeeld CAT, PICO), verslaglegging, brieven, statusnabespreking, onaangekondigde simulatiepatiënten, praktijk audits, voorschrijfgedrag, patiëntuitkomsten
SHOWS HOW	Stationstoetsen, simulatietoetsen, toetsen met Standardized Patients
KNOWS HOW	Casusgebonden schriftelijke toetsen (Multiple choice, essays, key-feature, etc.), mondelinge examens gericht op klinische toepassing
KNOWS	Schriftelijke toetsen (Multiple choice, essays, key-feature, etc.), mondelinge examens gericht op feitenkennis

enkele praktijkopdrachten of -besprekingen en een 360 graden feedbackrapportage in zijn/haar portfolio opneemt.

Beslissingen over bekwaamheid respectievelijk certificering worden dus altijd genomen op basis van geaggregeerde informatie over het functioneren van de aios, verzameld over een langere tijdsperiode en afkomstig van verschillende informatiebronnen. *Eén enkele beoordeling (werkplekbeoordeling c.q. toetsresultaat) mag nooit de enige en unieke basis vormen voor het nemen van summatieve beslissingen!* Het proces van informatieaggregatie en besluitvorming is vergelijkbaar met de besluitvorming bij gezondheidsklachten van patiënten: één enkel symptoom of signaal is onvoldoende om beleid te kunnen uitstippelen. Definitief beleid wordt pas vastgesteld na zorgvuldige selectie, ordening en interpretatie van relevante gegevens uit verschillende informatiebronnen in het patiëntendossier (zoals anamnese, lichamelijk onderzoek, laboratoriumonderzoek en ander aanvullend onderzoek), zonodig na consultatie van collega's.

Interpretatie en gebruik van toetsresultaten in begeleidingsgesprekken

Het gebruik van het portfolio, hetzij bij het coachen of sturen van het leerproces, hetzij bij het nemen van beslissingen over voortgang en bekwaamheid, veronderstelt dat de in het portfolio verzamelde informatie een representatief beeld geeft van het feitelijke functioneren van de aios in de opleiding. Toetsresultaten vormen een deel van de relevante informatie. Prescriptieve richtlijnen voor wat betreft noodzakelijke aantallen en typen beoordelingen voor het nemen van beslissingen c.q. doen van uitspraken zijn niet beschikbaar. Dit hangt sterk af van het competentiedomein of de taak waarover we uitspraken willen doen, maar ook van het individuele leertraject van de aios. Afhankelijk van de wijze waarop de aios functioneert en afhankelijk van het risico dat verbonden is aan de te beoordelen activiteiten, kan behoefte bestaan aan meer of minder (formele) beoordelingsmomenten. Patiëntveiligheid is daarbij een belangrijk criterium. Beslissingen vragen daarom steeds om een verstandige interpretatie van de informatie in het portfolio, gegeven de

actuele opleidingssituatie van de aios en de eisen in het opleidingsplan.

Bij interpretatie en gebruik van toetsresultaten is een aantal aandachtspunten van belang:

- *Maak samenvattende overzichten:* Om het leren van de aios te coachen en te sturen is het noodzakelijk inzicht te hebben in sterke kanten en in mogelijke verbeterpunten in het functioneren, bijvoorbeeld op basis van samenvattende overzichten per competentiedomein (algemeen en/of specialisatiegebonden). Elektronische portfoliosystemen kunnen hierbij ondersteunen door het genereren van overzichtsrapportages, al dan niet uitgesplitst naar (deel)competenties en opleidingsthema's. Vooral de *geschreven* (narratieve) feedback geeft concrete handvatten voor gerichte sturing van het leerproces. Gemiddeld behaalde scores hebben vooral een signaleringsfunctie.
- *Check scores en feedback die afwijken van de trend:* Op basis van één enkele werkplekbeoordeling wordt specifieke informatie verkregen over het functioneren op dat moment, ten aanzien van één specifieke taak of patiëntenprobleem (bijvoorbeeld jong kind met buikklachten) in één specifieke situatie (bijvoorbeeld poli, afdeling, SEH). Feedback en behaalde scores op separate (werkplek)beoordelingen (met name uitschieters naar boven en beneden) kunnen daarmee een gedetailleerde indicatie geven van specifieke bekwaamheden en aandachtspunten in het functioneren. Discrepancies tussen zelfoordelen en oordelen van anderen (bijvoorbeeld in de 360 graden feedback) kunnen een belangrijk hulpmiddel zijn bij het stimuleren van reflectie.
- *Baseer beslissingen uitsluitend op geaggregeerde en gevarieerde informatie:* Uit-

spraken over *competenties* zijn niet mogelijk op basis van één enkele toets of toetsinstrument. Geen enkele toets of toetsinstrument voldoet aan alle kwaliteitseisen wat betreft betrouwbaarheid en validiteit. Uitspraken over competenties zijn daarom bij voorkeur gebaseerd op een variatie aan toetsresultaten, afkomstig van verschillende instrumenten en bronnen. In het portfolio dienen in elk geval beoordelingen op het 'DOES'-niveau (Miller) aanwezig te zijn, met een zekere regelmaat en gespreid in de tijd afgenomen. Let daarnaast op variatie in werkplekbeoordelingen: is er sprake van voldoende variatie in patiëntproblematiek? Voldoende variatie in beoordelingssituaties? Voldoende variatie in beoordelaars? In het algemeen geldt dat veel beoordelingen noodzakelijk zijn om te kunnen komen tot betrouwbare uitspraken over het functioneren in een bepaald (competentie)domein. Voor werkplekbeoordelingen die gebaseerd zijn op individuele patiëntcasus (bijvoorbeeld KPB, OSATS) geldt dat 6-11 beoordelingen noodzakelijk zijn om te komen tot betrouwbare oordelen. Wat betreft het aantal beoordelaars geldt ten aanzien van werkplekbeoordelingen dat 6-10 beoordelaars nodig zijn om van betrouwbare oordelen te kunnen spreken (grotere aantallen als het gaat om patiëntoordelen!).

- *Geschreven (narratieve) feedback is belangrijker dan scores:* Bij de interpretatie van toetsresultaten, en in het bijzonder van werkplekbeoordelingen, zijn behaalde scores minder informatief dan geschreven feedback. Interpretatie van scores kan lastig zijn als er gebruik gemaakt wordt van uiteenlopende toetsinstrumenten met verschillende scoringsmodellen. Daarnaast is er vaak sprake van milde oor-

delen (score-inflatie) en kunnen beoordelaars van elkaar verschillen wat betreft de manier waarop toetsinstrumenten gebruikt worden (verschillen in gehanteerde referentiekaders, criteria, e.d.). Het is daarom belangrijk dat goede afspraken gemaakt worden in opleiderteams over het gebruik van instrumenten.

- *Trianguleer de informatie:* Toetsresultaten zijn niet de enige bron van informatie. Trianguleer de verkregen informatie, dat wil zeggen: vergelijk toetsresultaten met elkaar en met andere informatiebronnen, waaronder het expertoordeel van het opleiderteam. Wijst alle informatie in dezelfde richting? Is er (duidelijk) sprake van groei? Bij twijfel of bij niet consistente informatie moet aanvullende informatie verzameld worden (in de vorm van werkplekbeoordelingen, observaties, of anderszins).
- *Neem beslissingen op basis van verdedigbare standaarden:* In opleidingsplannen worden minimumeisen (absolute standaarden) beschreven waaraan aios op verschillende momenten in de opleiding dienen te voldoen. Er is echter nog maar beperkt evidence-based informatie beschikbaar over leertrajecten van aios in de medisch-specialistische vervolgopleidingen, zeker als het gaat om de algemene competenties. Daarom wordt aanbevolen om scores en competentieontwikkeling van de individuele aios steeds te vergelijken met die van een adequaat gekozen referentiegroep, bijvoorbeeld aios uit dezelfde opleidingsregio in een zelfde fase van opleiding. Ook hierbij kan ondersteuning geboden worden door elektronische portfoliosystemen. Maak in ieder geval in het opleiderteam goede afspraken over te hanteren criteria en standaarden (organiseer consensusbesprekingen).
- *Motiveer en documenteer beslissingen:* Maak duidelijke afspraken over besluitvormingsprocedures. Pas procedures aan aan de zwaarte van de beslissing: gebruik meerdere beoordelaars en consensusprocedures als beslissingen lastig en zwaarwegend zijn. Neem beslissingen over de voortgang in de opleiding als team. Geef de aios, waar mogelijk, een stem in de besluitvorming. Motiveer en documenteer summatieve beslissingen (bekwaamverklaringen, voortgangsbepalingen).

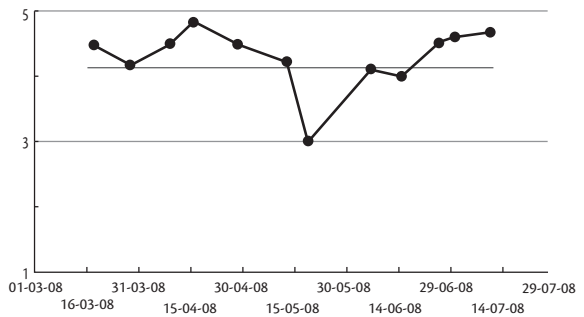
Randvoorwaarden

Beschreven gebruik van toetsen, waarbij toetsresultaten zowel gebruikt worden om het leerproces te sturen als om beslissingen te onderbouwen, vereist een veilige omgeving voor alle betrokkenen. Maak duidelijke afspraken met alle betrokkenen (aios en staffleden) over de wijze waarop toetsing en begeleiding plaatsvinden; helderheid over wederzijdse verwachtingen, criteria, standaarden en procedures draagt bij aan een veilig leerklimaat. Veiligheid neemt toe naarmate evaluatie en documentatie van functioneren meer ingebed worden in dagelijkse werkrouines. Draag als opleiderteam gezamenlijk verantwoordelijkheid voor de kwaliteit van toetsing. Zorg – waar nodig – voor professionalisering van staf en aios.

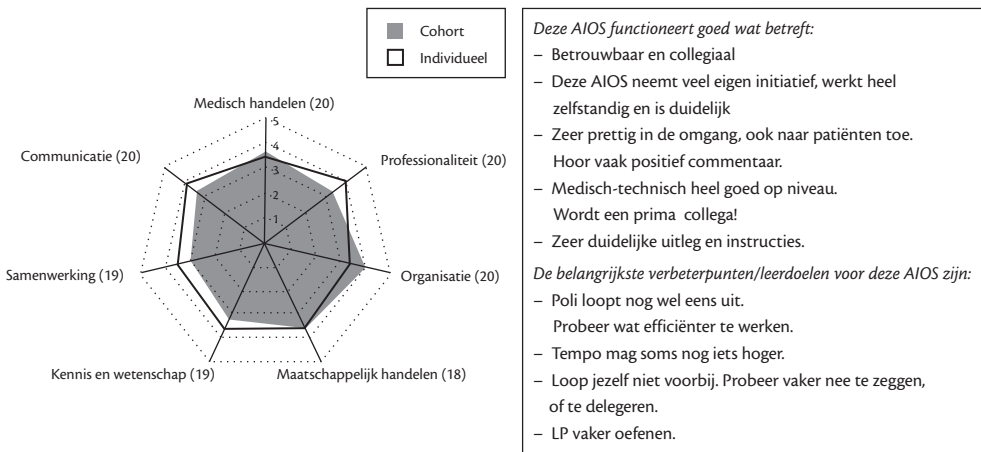
Voorbeelden van een geaggregeerde rapportage over een aios

In de onderstaande drie figuren zijn enkele voorbeelden afgebeeld van aggregatie van toetsresultaten uit aios-portfolio's. Elektronische portfoliosystemen kunnen ondersteunen bij het maken van dergelijke samenvattende overzichten. De voorbeelden in de Figuren 2, 3 en 4 zijn afkomstig van MMTSS/ePASS (ontwikkeld in OOR Zuid-Oost Nederland).

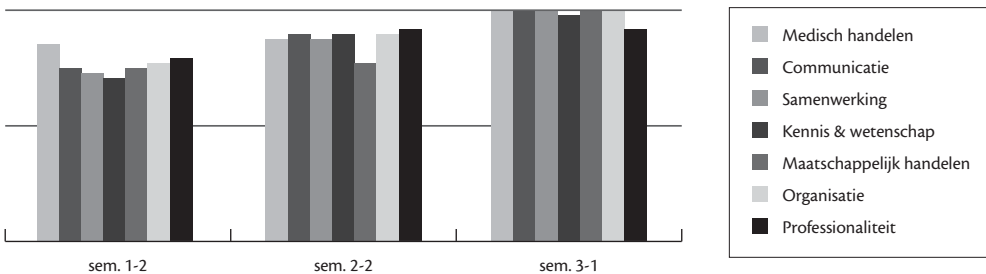
Overzicht beoordelingen	Datum	Score Communicatie
KPB poli algemeen	2008-03-18	4,5
Referaat	2008-03-29	4,2
...		
KPB slecht nieuws	2008-05-19	3,0
360-graden feedback	2008-06-06	4,0
...		
KPB poli algemeen	2006-06-26	4,5
...		



Figuur 2. Overzicht van toetsmomenten over een periode van enkele maanden en van behaalde scores voor het competentiedomein ‘Communicatie’, per toetsmoment.



Figuur 3. Gemiddeld behaalde scores per CanMEDS domein in één semester, berekend over alle toetsinstrumenten, en bijbehorend overzicht van narratieve feedback op beoordelingsformulieren. Individueel behaalde scores worden vergeleken met gemiddeld behaalde scores van een referentiegroep (aios in het zelfde opleidingscohort).



Figuur 4. Gemiddeld behaalde scores per CanMEDS domein, berekend over alle toetsinstrumenten, voor verschillende fasen (semesters) in de opleiding.