

Achtergronden bij de Richtlijn multisource feedback voor de aios

1. Achtergrond multisource feedback

De medisch specialist van de 21ste eeuw wordt geacht over verschillende competenties te beschikken, zoals bijvoorbeeld opgesteld door de Accreditation Council for Graduate Medical Education (ACGME),⁴⁾ Royal College of Physicians and Surgeons of Canada CanMEDS roles,⁵⁾ of United Kingdom's Best Medical Practice framework.⁶⁾ In Nederland wordt in de medische vervolgoopleidingen gebruik gemaakt van de competenties die gebaseerd zijn op de CanMEDS rollen. De noodzaak om specialisten in opleiding niet alleen op hun kennis maar op alle competentiegebieden te beoordelen, heeft geleid tot de ontwikkeling van nieuwe beoordelingsmethoden, waaronder MSF. Bij deze vorm van beoordeling, in de literatuur ook wel 360 graden beoordeling genoemd, wordt zowel aan personen uit de werkomgeving als aan patiënten gevraagd feedback te geven op het functioneren van een arts die zij direct en herhaaldelijk hebben kunnen observeren. MSF is bruikbaar voor het beoordelen van competenties als professionaliteit, samenwerking en communicatie. Deze competenties kunnen niet uitsluitend door een eenmalige observatie of op een traditionele wijze worden getoetst; feedback van collegae en patiënten is onmisbaar voor een volledige beoordeling¹⁻².

MSF bestaat uit specifieke processen en instrumenten voor informatieverzame-

ling, beoordeling en feedback op de werkvloer. Vragenlijsten, ontworpen om gegevens over specifieke gedragsaspecten (bijvoorbeeld communicatievaardigheden, professionaliteit, teamwork) te verzamelen, worden door verschillende respondenten ingevuld over één persoon. MSF is ontstaan in het bedrijfsleven, waar de werknemers niet vaak of niet van dichtbij genoeg door de supervisor werden geobserveerd om zinvolle feedback te kunnen ontvangen. Collegae op de werkvloer waren vaak in een betere positie om feedback over de werknemer te verstrekken. Door verschillende collegae vragenlijsten te laten invullen kon de informatie van meerdere medewerkers worden samengevoegd en gebruikt om feedback te verstrekken. In het bedrijfsleven is MSF geaccepteerd omdat informatie van meerdere individuen (waarnemers) en veelvoudige beoordelaars (groepen waarnemers) wordt gebruikt en waarbij de bias die inherent is aan de beoordeling van 'één persoon, één perspectief' (dat wil zeggen, het perspectief van de supervisor) wordt verminderd.¹⁻².

In medische settings kunnen de waarnemers (ook wel beoordelaars, respondenten of feedbackgevers genoemd) een verschillende achtergrond hebben. Te denken valt aan medische collegae (bijvoorbeeld peers – aios in vergelijkbare opleidingsfase –, supervisor, verwijzende artsen, coassisten-

4) ACGME Competencies: <http://www.acgme.org/outcome/comp/compFull.asp>.

5) RCPSC CanMEDS Competencies: <http://rcpsc.medical.org/canmeds/index.php>.

6) GMC (UK) Good Medical Practice: http://www.gmc-uk.org/guidance/good_medical_practice/index.asp.

ten), niet-medische collega's (bijvoorbeeld verpleegkundigen, paramedici en psychologen) en patiënten en hun familieleden. Deze uit diverse (beroeps)groepen afkomstige waarnemers hebben verschillende perspectieven en observaties betreffende het handelen van artsen, afhankelijk van de context waarin zij samenwerken met de arts. Bovendien kunnen zij het functioneren van een arts vaak gedurende langere tijd observeren.

MSF kan worden gebruikt in het gehele continuüm van medisch onderwijs, hoewel het grootste deel van de gepubliceerde studies zich op aios³⁻⁷ en praktiserend specialisten richt.⁸⁻¹³ In deze verantwoording wordt ingegaan op de belangrijkste aspecten van MSF en worden de in de Richtlijn gegeven adviezen, mede aan de hand van relevante literatuur, toegelicht.

2. Doel van de multisource feedback-procedure

De beoordeling aan de hand van gestructureerde vragenlijsten heeft als doel het verkrijgen van informatie over het functioneren van de aios op een aantal competentiegebieden, zodat hierover zinvolle feedback kan worden gegeven. In de literatuur worden zowel summatieve als formatieve doelen van MSF beschreven. Bij een summatief doel weegt de beoordeling mee voor de beslissing van het wel of niet voortzetten van de opleiding. De procedure van MSF, zoals geadviseerd in deze Richtlijn, heeft primair een formatief doel: de beoordeling zal gericht zijn op het aanzetten tot reflectie, het opstellen van doelen en het verbeteren van het functioneren. De resultaten zullen in voortgangsgesprekken worden besproken met de aios (zie hoofdstuk 13). Het secundaire doel van de procedure zal wel summatief van aard zijn: het signaleren van 'problematisch functionerende artsen'. Wanneer een aios een onvoldoende beoordeling krijgt op meer-

dere competentiegebieden kan er bijvoorbeeld voor gekozen worden de aios intensiever te begeleiden en hem/haar vaker de MSF procedure te laten doorlopen.

Gezondheidszorgorganisaties die een vorm van MSF hebben ontwikkeld en ingevoerd kunnen verschillende doelstellingen hebben, bijvoorbeeld de kwaliteitsverbetering van hele groepen, maar ook verbetering van het functioneren op het individuele niveau. In Alberta, Canada, bijvoorbeeld, wilde het Royal College of Physicians and Surgeons of Alberta een formatief programma voor kwaliteitsverbetering dat als eerste screening van het niveau van alle artsen zou dienen.¹⁴ In het Verenigd Koninkrijk wilde de General Medical Council een beoordeling voor specialisten in opleiding, gebaseerd op hun omschrijving van Best Medical Practice. Zowel de Sheffield Peer Review Assessment Tool (SPRAT)³⁻⁴ als het Pre-registration House Officers Appraisal and Assessment System (PHAST)⁵ zijn ontwikkeld om artsen te identificeren bij wie verbetering van het functioneren nodig is. Mini-PAT (een kortere versie van SPRAT) wordt gebruikt voor basisartsen in de eerste twee ('foundation') jaren van hun algemene specialisatie. Hoewel het instrument bedoeld is om artsen te identificeren die ondermaats presteren, maakt het deel uit van een groter portfolio dat ontworpen is om aios te helpen leerbehoeften en doelen vast te stellen en een actieplan te maken.⁶ Evans et al. suggereren dat zelfs wanneer de MSF een formatief doel heeft, het ook kan worden gebruikt om mogelijke 'probleemgevallen' te signaleren.¹⁵ Ook bij formatieve MSF procedures bestaat er een ethische verplichting tegenover patiënten om hen te beschermen tegen deze uitzonderingen en zo nodig actie te ondernemen.¹

Het is belangrijk het doel van de MSF procedure uit te leggen aan de beoorde-

laars die ermee moeten werken. Te verwachten valt dat beoordelaars vragenlijsten anders invullen en aios mogelijk andere beoordelaars selecteren bij het gebruik van de beoordeling voor zwaarwegende beslissingen, dan bij het gebruik voor reflectie en gedragsverbetering. Beoordelaars zullen milder zijn in hun feedback wanneer zij weten dat er veel van de beoordeling afhangt.¹⁶ Wanneer zij er niet op vertrouwen dat de beoordeling primair voor een formatief doeleinde wordt gebruikt, zullen zij waarschijnlijk ook minder eerlijk zijn in hun geschreven opmerkingen.¹ Door beoordelaars op de hoogte te brengen van het doel van de MSF zal de motivatie tot invullen mogelijk toenemen. Bovendien kan het van belang zijn beoordelaars op de hoogte te stellen van het aantal andere beoordelaars. Op de website *www.multisource-feedback.nl*, die wordt gebruikt voor de MSF procedure, voorgesteld in deze Richtlijn, staat informatie voor de beoordelaars. Deze informatie kan door de aios worden uitgeprint en uitgedeeld of per e-mail worden verstuurd. Tevens wordt bij de vragenlijsten voor zowel patiënten als collegae uitgelegd wat het doel van de procedure is en hoe de gegevens worden verwerkt en teruggekoppeld naar de aios.

3. Geschikte beoordelaars voor multisource feedback

In bijlage I wordt een overzicht gegeven van gebruikte beoordelaars in de tot heden uitgevoerde studies over MSF in een geneeskundige setting.⁷⁾ In vrijwel alle studies wordt een onderscheid gemaakt tussen medische collegae en andere werkers in de gezondheidszorg; ook in deze Richtlijn is voor deze indeling gekozen. In Tabel

2 worden per groep voorbeelden gegeven van mogelijke beoordelaars die samenwerken met de aios en die vanuit hun perspectief een oordeel kunnen vormen over het functioneren van de aios gedurende een langere periode en in verschillende situaties.

Tot nu toe hebben slechts weinig studies gebruik gemaakt van beoordelaars uit deze vier groepen. Zelfbeoordeling wordt vaak beschouwd als een van de bronnen bij 360 graden beoordelingen; vanwege de andere positie in het spectrum van informatie over het functioneren wordt deze categorie vooralsnog buiten beschouwing gelaten. Een toelichting volgt in hoofdstuk 8.

Met name patiënten worden lang niet altijd betrokken in de beoordeling. Hoewel beoordelingen van de prestaties van artsen door patiënten in principe waardevol zijn, leveren ze ook een aantal problemen op. Ten eerste zijn grote aantallen beoordelingen nodig; in sommige studies wordt zelfs gesproken van 50 om een voldoende betrouwbaarheid te verkrijgen.⁷ Ernstig zieke patiënten vullen de formulieren vaak niet in, en diegenen die wel invullen geven vaak een lagere beoordeling dan matig zieke patiënten. Verder kunnen patiënten niet altijd goed onderscheid maken tussen verschillende aspecten van klinisch handelen en geven ze vaak hoge beoordelingen. Deze beperkingen maken het lastig de patiënt als enige bron te gebruiken.¹⁷ Toch kan het perspectief van patiënten in dit tijdperk van patiëntgerichte zorg niet meer buiten beschouwing worden gelaten. Patiënten, de doelgroep van de zorg en het medisch handelen, vormen een essentiële bron voor het beoordelen van de omgang met patiënten door aios. Problemen of conflic-

7) Met bronnen en beoordelaars wordt hetzelfde bedoeld, voor de leesbaarheid worden beide termen afwisselend gebruikt.

Tabel 2. *Categorieën beoordelaars.*⁸⁾

Medische collegae	Andere werkers in de zorg	Patiënten	Zelfbeoordeling
Supervisor	Verpleegkundigen (aan het bed, team-leider, hoofd, gespecialiseerd, SEH)	Patiënten en kinderen ouder dan 15 jaar	(zie hoofdstuk 8)
Staflid (dienst)		Ouders van patiëntjes jonger dan 16 jaar	
Opleider	Verloskundige	Begeleiders bij patiënten die zelf de vragenlijsten niet kunnen invullen	
Coassistent	Physician assistant		
Aios (peers)	Nurse practitioner		
Aios (senior)	Operatieassistent		
Aios van een aanpalend specialisme	Paramedicus		
Anios	Maatschappelijk werker		
Consulent	Psycholoog		
Huisarts/ verwijzend arts	Secretaresse		

ten tussen artsen en patiënten doen zich vaak voor op het gebied van communicatie en professionaliteit. Daarnaast blijkt de bereidheid tot medewerking aan het geven van feedback aan artsen groot te zijn.^{7 9 12-13}

In géén van de gevonden studies werd aan de partners van patiënten gevraagd om het functioneren van artsen te beoordelen. In studies op het gebied van de Kindergeneeskunde worden wel ouders van patiënten ondervraagd, er wordt echter niet beschreven bij welke leeftijdscategorie van patiënten.¹¹ Kinderen in de leeftijd

van 12 tot 16 jaar hebben juridische zeggenschap volgens de Wet op de Geneeskundige Behandelingsovereenkomst (WGBO), maar worden in de bestaande MSF procedures niet gevraagd om hun mening. Crossley et al. verrichtten in het Verenigd Koninkrijk een studie naar de haalbaarheid van de beoordeling van het functioneren van kinderartsen door kinderen en/of hun ouders.¹⁸ Zij ontwikkelden verschillende vragenlijsten over de niet-medische aspecten van het functioneren van kinderartsen, geschikt gemaakt voor jonge kinderen (7-11 jr.), oudere kinderen (11-16 jr.) en vol-

8) Voorbeelden van beoordelaars o.b.v. literatuur en uitkomsten van InVIVO Focus-bijeenkomst dd. 16-04-2007
Verslag focusbijeenkomst: <http://www.medischevervolgopleidingen.nl/content/documenten/invivo/1.pdf>.
www.medischevervolgopleidingen.nl → implementatie → verslag open-space sessies.

wassenen (> 16 jr.). Bepaling van de haalbaarheid bestond onder andere uit het bepalen van het benodigd aantal beoordelaars voor een betrouwbaar oordeel. Uit de analyses van de beoordelingen van 352 arts-patiënt interacties bleek dat 15 beoordelaars ouder dan 16 jaar voldoende zijn voor een betrouwbaar oordeel. Bij beoordeling door kinderen jonger dan 16 jaar zouden meer dan 30 beoordelaars nodig zijn. Aangezien ouders vrijwel altijd aanwezig zijn bij het contact tussen kinderarts en patiënt, werd geconcludeerd dat ouders een geverifieerde en betrouwbare bron zijn voor het beoordelen van de communicatie over en weer. Wij adviseren in de Kindergeneeskunde kinderen ouder dan 16 jaar en ouders en/of verzorgers van kinderen tot 16 jaar te kiezen als beoordelaars. Mutatis mutandis kan dat ook gelden voor begeleiders van patiënten die om andere redenen de vragenlijst niet kunnen invullen.

Het meeste onderzoek is verricht naar de validiteit en betrouwbaarheid van de beoordeling door collegae van dezelfde discipline of uit hetzelfde opleidingsjaar, peer assessment genoemd.^{3-4 6 15 19-20} Van beoordelingen door peers is aangetoond dat ze consistent zijn, onafhankelijk van de selectiemethode van beoordelaars. De beoordelingen zijn wel afhankelijk van vertrouwen en vereisen een grote aandacht voor veiligheid van gegevens, anders kunnen ze een ondermijnende, destructieve en verdelende werking hebben.¹⁷ Voor beoordelingen door opleiders en stafleden die supervisie bieden aan aios worden vergelijkbare resultaten beschreven.²¹ In de groep medische collegae blijken consulenten en verwijzende artsen vaak minder goed zicht te hebben op het functioneren van aios.¹²

Wat betreft andere werkers in de zorg worden verpleegkundigen, verloskundigen, psychologen en paramedici beschreven in de literatuur. Vooral in de eerste

procedures voor MSF die zijn ontwikkeld, werden verpleegkundigen nog maar weinig betrokken. Eén van de redenen hiervoor was dat artsen (in opleiding) moeie zouden hebben verpleegkundigen te accepteren als geloofwaardige bron voor een gefundeerde evaluatie van hun prestaties.²² Beoordelingen door verpleegkundigen kunnen echter waardevol zijn gezien de nauwe samenwerking op vooral verpleegafdelingen, en zijn voldoende betrouwbaar bij zes tot tien personen.²³⁻²⁴ De beoordelingen blijken goed te correleren met de beoordeling door patiënten en opleiders op het gebied van interpersoonlijke vaardigheden.⁷ Uit meerdere studies blijkt dat medische secretaresses geen geschikte bron te zijn voor het beoordelen van het functioneren van aios, aangezien zij de arts niet meemaken in de spreekkamer en niet direct het contact met patiënten kunnen observeren.

Geadviseerd wordt de volgende beoordelaars te laten deelnemen aan de MSF:

- *Medische collegae*: supervisor, staflid (dienst), opleider, consulent, coassistent, aios (peers), aios in hogere opleidingsjaren, anios, huisarts/verwijzend arts.
- *Andere werkers in de zorg*: verpleegkundige (aan het bed), verpleegkundig teamleider/manager, gespecialiseerd verpleegkundige, verloskundige, physician assistant, nurse practitioner, paramedicus (fysiotherapeut, diëtist), psycholoog, pedagogisch medewerker, maatschappelijk werker.
- *Patiënten*: liefst na meerdere contacten met de aios en alleen in bepaalde settings (op IC bijvoorbeeld minder haalbaar). Bij Kindergeneeskunde: kinderen vanaf 16 jaar en ouders van patiënten < 16 jaar.

In hoofdstuk 6 wordt nader ingegaan op bovenstaande criteria voor het selecteren van geschikte beoordelaars.

4. Gewenst aantal beoordelaars

Het vereiste aantal beoordelingen per groep beoordelaars voor het verkrijgen van voldoende betrouwbare en valide gegevens is internationaal uitgebreid onderzocht. Het benodigde aantal beoordelingen is over het algemeen gebaseerd op de haalbaarheid, het doel van de beoordeling (formatief versus summatief) en op de psychometrische kenmerken die gewenst zijn.¹ Wood et al. beschrijven dat het benodigde aantal beoordelaars deels gerelateerd is aan de mogelijkheid van beoordelingsfouten of mate van bias. Het gebruik van beoordelingen door meerdere beoordelaars vermindert de mogelijkheid van bias voor een deel vanzelf, omdat men niet afhankelijk is van één enkele beoordeling. De informatierijkheid van de feedback zal toenemen naarmate meerdere soorten bronnen worden gevraagd om feedback. De vraag blijft echter hoeveel beoordelaars per peiling nodig zijn om de kans op een onjuiste, niet representatieve beoordeling te minimaliseren.²⁵ In veel MSF studies is dan ook de generaliseerbaarheid van vragenlijsten onderzocht. Hierbij maakt men een schatting van de betrouwbaarheid van verschillende combinaties van aantallen beoordelaars en aantallen vragen door het aantal beoordelaars te variëren. Opzet van het merendeel van deze studies was het aantonen van het benodigde aantal beoordelaars voor het behalen van een generaliseerbaarheidscoëfficiënt (E_p^2) van 0.7 of meer.^{7 10-13 20 26} Deze waarde wordt in zowel het bedrijfsleven als in de medische settings als voldoende gezien en houdt in dat er 70% kans is dat verschillen in scores een daadwerkelijk verschil tussen artsen aantonen, in plaats van verschillen tussen beoordelaars. Of anders gezegd: er is 70% kans op reproduceerbare gegevens bij een herhaalde meting. Voor de volledigheid wordt een aantal van deze studies hier genoemd. Ramsey et al. analyseerden

3005 vragenlijsten met betrekking tot MSF van 187 artsen. Tien tot elf reacties van collega-artsen waren noodzakelijk om een generaliseerbaarheidcoëfficiënt van 0.7 te bereiken.²⁰ Bij beoordeling van dezelfde groep, maar dan door verpleegkundigen, bleek een aantal van 1015 nodig om voldoende generaliseerbare gegevens op te leveren.²² Butterfield & Mazzaferrri schatten in 1991 dat slechts vijf of zes verpleegkundigen een representatief beeld van elke aios konden verstrekken.²⁷ Woolliscroft et al. (1994) hadden meer teleurstellende resultaten. Zij concludeerden dat vijf tot tien stafleden nodig zouden zijn voor een representatieve beoordeling van aios op interpersoonlijke en professionele kwaliteiten. Het benodigde aantal verpleegkundigen was echter 10 tot 20, en voor patiënten bedroeg het aantal zelfs meer dan 50.⁷ Ramsey en Wenrich suggereren in 1999 dat totaal zeven tot twaalf beoordelaars nodig zijn voor betrouwbare feedback over prestaties.²⁸ Dit aantal varieert afhankelijk van het te toetsen domein of competentiegebied; bij inter-persoonlijke vaardigheden en de competentie professionaliteit zijn bijvoorbeeld meer beoordelaars nodig.²⁵ Voor de MSF in het kader van het Canadese herregistratiesysteem voor medisch specialisten, Physician Achievement Review, wordt gebruik gemaakt van tien medische collega's, tien andere werkers in de zorg en 25 patiënten.^{9 13-14} Zie Bijlage I van deze Richtlijn voor het aantal benodigde beoordelingen per groep beoordelaars (een selectie van verrichte studies naar MSF).

Generaliseerbaarheidstudies zijn verricht om het aantal beoordelaars te bepalen dat nodig is om op betrouwbare wijze problematisch functionerende artsen te kunnen signaleren, ofwel te gebruiken als summatieve beoordelingsmethode. Davies et al. toonden in 2005 op basis van de generaliseerbaarheidstheorie aan dat bij

gebruik van SPRAT voor kinderartsen in opleiding, vier beoordelaars die 24 vragen beantwoorden voldoende zijn om vast te kunnen stellen of artsen wel of niet slecht functioneren. Dit is volgens de auteurs vooral praktisch in bijvoorbeeld de Huisartsgeneeskunde, waar vaak weinig beoordelaars beschikbaar zijn. Zij geven wel aan dat voor het aantonen van 'grensgeslagen' meer beoordelaars nodig zijn. Over het algemeen geldt dat wanneer een summatief doel wordt nagestreefd er meer beoordelaars nodig zijn dan bij een formatief doel van de procedure.⁴

De gevonden minimaal benodigde aantallen beoordelaars lopen sterk uiteen tussen de verschillende studies, afhankelijk van de inhoud, het doel en de lengte van de vragenlijst alsmede van de setting en mate van training van beoordelaars. Een andere reden voor de verschillen in aantallen die zijn gevonden in onderstaande studies is dat in sommige groepen een tendens kan bestaan om inconsistent te beoordelen. Beoordelaars kunnen een verschillende opleiding, achtergrond en beroepsnormen hebben. Tevens kunnen zij uiteenlopende ervaringen hebben met de aios, waardoor de mate van overeenstemming tussen scores van verschillende beoordelaars niet hoog kan zijn. Ten derde kan het vragen om een beoordeling over een te groot aantal categorieën mogelijk leiden tot een grotere subjectiviteit en irrelevantie en vaker tot een onvermogen van de beoordelaar om bepaalde categorieën te beoordelen. Tot slot kan er ook nog een verschil zijn in training ten aanzien van de MSF procedure. Deze factoren suggereren dat wanneer er valide, meer gerichte beoordelingsinstrumenten worden ontworpen en beoordelaars worden getraind, het aantal benodigde beoordelaars per peiling mogelijk gereduceerd kan worden.²⁵

Op basis van de wetenschappelijke literatuur en resultaten van de in Nederland verrichte pilotstudie, worden in deze Richtlijn de volgende aantallen per groep beoordelaars geadviseerd (in hoofdstuk 4 zijn deze aantallen toegelicht; zie ook hoofdstuk 11C):

- *Medische collegae:* totaal minimaal zes beoordelaars
 - *Andere werkers in de zorg:* totaal minimaal zes
 - *Patiënten:* totaal minimaal tien beoordelaars
- } Totaal minimaal 22 beoordelaars

Aangenomen mag worden dat de betrouwbaarheid bij deze aantallen beoordelaars voldoende is voor een summatief oordeel over de aios. Als de MSF procedure uitsluitend een formatief doel dient, kan wellicht met minder beoordelaars worden volstaan. Dit zal voorwerp van nader onderzoek moeten zijn.

5. Te beoordelen competenties door verschillende beoordelaars

In de literatuur worden gewenste gedragingen van artsen op verschillende manieren beschreven, bijvoorbeeld als gedragsdomeinen, competentiegebieden of kritische gedragingen.^{3 5-6 10-11 29} De inhoud van deze verschillende beschrijvingen komt grotendeels overeen met de inhoud van de CanMEDS competentiegebieden, die zijn vertaald naar de Nederlandse praktijk.³⁰ De volgende onderwerpen voor MSF worden beschreven in de literatuur: medisch handelen, management van patiëntenzorg, interpersoonlijke kwaliteiten, communicatieve vaardigheden, professionaliteit, collegialiteit/samenwerking, organiseren, medische praktijkvoering, onderhouden van de professionele ontwikkeling, (telefonische) bereikbaarheid. Maatschappelijk handelen

wordt niet als zodanig beschreven in de literatuur; wel wordt in verschillende studies gevraagd naar omgang met incidenten in de zorg, mate van preventieve zorg en naar handelen volgens relevante wetgeving.

Verskillende beoordelaars kunnen verschillende aspecten van het functioneren van artsen beoordelen en deels bestaat er een overlap tussen de competenties die zij kunnen beoordelen. Idealiter wordt beoordelaars alleen gevraagd gedrag te beoordelen dat zij daadwerkelijk in de praktijk en liefst herhaaldelijk hebben kunnen observeren.¹ Zo zal een beoordelaar die de arts nooit heeft meegemaakt in de spreekkamer of aan het bed van de patiënt niet in staat zijn de communicatie met patiënten en familie te beoordelen. De patiënt is echter bij uitstek een geschikte bron om over dit punt een oordeel te geven, evenals verpleegkundigen die bij gesprekken met patiënten aanwezig zijn. Om een keuze te maken welke beoordelaars geschikt zijn voor het beoordelen van bepaalde competenties van de aios, zal dus bepaald moeten worden welk gedrag een beoordelaar daadwerkelijk kan observeren en beoordelen. Daarbij spelen ook haalbaarheid en acceptabiliteit van het ondervragen van mensen met een verschillende (beroeps)achtergrond een belangrijke rol. Het zal bijvoorbeeld lastig zijn vragenlijsten te laten invullen door patiënten met spoedeisende of psychiatrische aandoeningen, alsmede door ernstig zieke of terminale patiënten. Tevens zullen analfabeten en patiënten met te weinig begrip en/of kennis van de taal moeilijk te ondervragen zijn.¹

In de meeste artikelen wordt beschreven dat patiënten vooral gevraagd wordt naar interpersoonlijke kwaliteiten, communicatie en professionaliteit. Op de vragenlijsten, ontworpen in samenhang met deze Richtlijn, worden aan patiënten vragen gesteld over twee van de zeven com-

petentiegebieden. Om patiënten niet teveel te belasten en de haalbaarheid voldoende hoog te houden is het aantal vragen beperkt tot acht vragen over communicatie en acht vragen over professionaliteit. De auteurs veronderstellen dat patiënten over het algemeen onvoldoende zicht hebben op de competenties samenwerking en maatschappelijk handelen om hier een beoordeling over te kunnen geven. In de literatuur worden ook voorbeelden gegeven van vragenlijsten voor patiënten met vragen over organisatie, bijvoorbeeld medische praktijkvoering en (telefonische) bereikbaarheid.^{1-2 11-12 26} Op dit gebied zou gevraagd kunnen worden of de arts voldoende tijd besteedt aan de patiënt en hoe lang de patiënt gemiddeld moet wachten op de arts. Aangezien de beoordeling van dit soort vragen ook afhankelijk is van organisatorische factoren buiten de aios, zijn deze vragen achterwege gelaten. Getracht is alleen te vragen naar aspecten van het functioneren die geheel op de aios zijn terug te voeren.

Medische collegae en peers wordt in vrijwel alle studies gevraagd naar communicatie en professionaliteit en daarnaast naar medisch handelen, technische vaardigheden, samenwerken, organisatie, professionele ontwikkeling en gebruik van beoordelaars (kennis en wetenschap). Violato et al. beschrijven in 2006 dat vragen die collega-artsen moeilijk te beantwoorden vonden, waren gerelateerd aan professionele ontwikkeling en verbetering in de praktijk.¹¹ Verpleegkundigen en andere niet-medische professionals in de gezondheidszorg wordt gevraagd naar gegevens over de zorg voor patiënten, interprofessionele samenwerking en communicatie. Johnson en Cujec toonden een voldoende betrouwbaarheid aan van de beoordeling van deze competenties door verpleegkundigen op een intensive care afdeling.²⁴ Het wordt door artsen minder goed geaccep-

teerd wanneer andere werkers in de zorg het medisch handelen of technische vaardigheden beoordelen.²⁰⁻²² Op de vragenlijsten voor andere werkers in de zorg worden dan ook geen vragen gesteld over de competentie medisch handelen. In de eerste versie van de vragenlijst voor medische collegae werden hier ook geen vragen over gesteld.

Bij de evaluatie van de eerste pilotstudie bleek het echter onnatuurlijk om bij medische collegae vragen over medisch inhoudelijke kwaliteiten geheel achterwege te laten, omdat dit ook belangrijke feedbackpunten kan opleveren. In de tweede versie van de vragenlijst voor medische collegae zijn dan ook drie vragen over de competentie medisch handelen toegevoegd. Over communicatie en professionaliteit kunnen in principe alle beoordelaars zoals genoemd in hoofdstuk 3 vanuit hun positie een beoordeling geven. In Tabel 3 wordt een voorbeeld gegeven van verschillende vragen over deze competenties aan de verschillende groepen beoordelaars.

Het aantal vragen kan een rol spelen in de haalbaarheid van de procedure en voorkomen moet worden dat door te lange vragenlijsten de respons afneemt. Lockyer et al. tonen in hun studie aan dat het minimum aantal vragen 20 bedraagt voor alle vragenlijsten bij acht medische bronnen, acht niet-medische beoordelaars en 25 patiënten.²⁶ Hierbij dient wel opgemerkt te worden dat in deze studie de MSF een summatief doel had en er dus hogere eisen gesteld worden aan de psychometrische kenmerken van de vragenlijsten.

In Tabel 4 is per categorie beoordelaars weergegeven over welke competenties zij gevraagd worden feedback te geven in het programma *Multisourcefeedback.nl*. Omwille van de haalbaarheid, en uitgaand van een formatieve beoordeling, is het

aantal vragen beperkt gehouden tot 16 voor patiënten en respectievelijk 29 en 23 voor medische collegae en anderen. In hoofdstuk 11 worden de gemaakte keuzes voor het bepalen van de vorm en inhoud van de vragenlijsten nader toegelicht.

6. Selectie en benadering van beoordelaars

A. Selectie van de beoordelaars

Op basis van literatuurstudie en discussie in focusgroepen wordt geadviseerd om de beoordelaars niet door de supervisor of opleider, maar door de aios zelf te laten selecteren. Dit kan echter vragen oproepen over de betrouwbaarheid van gegevens: worden er geen 'vriendjes' uitgekozen, vullen collegae de vragenlijsten dan wel eerlijk in en blijft de veiligheid van relaties met patiënten en collegae wel gewaarborgd? Deze laatste vraag wordt in hoofdstuk 7 toegelicht. Het potentiële probleem van het alleen uitkiezen van 'vriendjes' wordt voor een deel opgelost door uit alle groepen meerdere beoordelaars te vragen. Het zal niet mogelijk zijn alleen maar collegae uit te kiezen met wie men ook buiten het werk bijvoorbeeld een vriendschappelijke relatie heeft. In een aantal studies is de mate waarin de beoordelaar en beoordeelde elkaar kennen, en de invloed hiervan op de scores, onderzocht. Deze factor blijkt verantwoordelijk te zijn voor minder dan 10-15% van de variantie van beoordelingen.¹⁸⁻²⁵⁻²⁶ Bovendien werd in een studie van Ramsey in 1993 aangetoond dat het selecteren van beoordelaars door de aios of selectie door een derde partij (supervisor/opleider) weinig verschil maakt voor de betrouwbaarheid van de beoordeling.²⁰ Hoe goed een beoordelaar de te beoordelen arts kent, en de invloed hiervan op de beoordeling, werd ook als onafhankelijke variabele onderzocht in een studie van Hall et al.⁹ Voor de peer assessmentvragenlijst toonden de resultaten een significante

Tabel 3. Voorbeelden van competenties, beoordelaars en vragen.

Competentiegebieden	Beoordelaars	Voorbeeldvragen/stellingen
Communicatie	Patiënten	De arts luisterde goed naar mij De arts gaf duidelijk uitleg
	Andere werkers in de zorg	De arts communiceert goed met patiënten en familie De arts geeft helder geformuleerde mondelinge opdrachten
	Medische collegae	De arts is zowel duidelijk als volledig in schriftelijke verslaglegging De arts presenteert tijdens besprekingen patiëntencasus accuraat en beknopt
Professionaliteit	Patiënten	Deze arts begreep mijn situatie goed (de situatie van mijn kind) Deze arts leek op de hoogte te zijn van mijn dossier/mijn situatie
	Andere werkers in de zorg	De arts stelt tijdens de dagelijkse werkzaamheden het belang van de patiënt centraal De arts kent het dossier en de toestand van de patiëntenzorg
	Medische collegae	De arts is zich bewust van eigen grenzen en handelt daarbinnen De arts accepteert feedback van anderen

Tabel 4. Overzicht van aantal vragen per competentiegebied en per categorie beoordelaars.

Bron	Aantal vragen	Competentiegebieden (aantal vragen)	Schaal
Medische collegae	29 + 1 open vraag*	Medisch handelen (3); communicatie (5), samenwerking (6), organisatie (4), kennis & wetenschap (3), professionaliteit (6), EPA's** (2)	5 waarden eens/oneens + categorie nvt
Anderen in de zorg	23 + 1 open vraag*	Communicatie (6), samenwerking (5), organisatie (4), kennis & wetenschap (2), professionaliteit (6)	
Patiënten	16 + 1 open vraag*	Communicatie (8), professionaliteit (8)	

* Open vraag: Heeft u tips en/of opmerkingen om het functioneren van deze arts te verbeteren?

** EPA: Deze aios vertrouw ik de volgende kritische activiteit toe om zelfstandig uit te voeren.³¹

invloed aan van de mate waarin de beoordelaar de arts kent. Voor bijna de helft van de vragen werden statistisch significante verschillen gevonden, waarbij beoordelaars die de arts niet of niet goed kenden een positievere beoordeling gaven dan beoordelaars die de artsen enigszins, goed of zeer goed kenden. Beoordelaars die de artsen niet goed kenden vormden slechts een klein percentage van alle respondenten (1.8%), en het gemiddelde verschil in scores was ook klein (0.21). Dezelfde analyse van de vragenlijst voor niet-medische collega's toonde overeenkomstige resultaten met opnieuw erg kleine verschillen in de scores. De vergelijking van zelfbeoordelingen met beoordelingen door peers, consultants en verwijzende artsen, en niet-medische collega's vertoonde weinig overeenstemming tussen de twee categorieën. De artsen beoordeelden zichzelf minder gunstig dan hun collega-professionals.

Naar aanleiding van de resultaten van bovengenoemde studies laat men in de meeste MSF procedures de selectie van beoordelaars over aan de te beoordelen arts zelf. Om wel een zo groot mogelijke spreiding van perspectieven te verkrijgen binnen een groep beoordelaars en te voorkomen dat veel vragen met 'niet te beoordelen' worden ingevuld, wordt een aantal voorwaarden voor de selectie van beoordelaars geadviseerd:

- *Medische collegae*: totaal minimaal zes beoordelaars; in ieder geval beoordeling door minimaal twee stafleden van het opleidingsteam (in ieder geval de huidige supervisor) en twee collega-aios waarmee een langere periode is samengewerkt. Overige beoordelaars zijn coassistenten, medische collegae met wie men dienst (weekend/week) heeft gedraaid, consultants en verwijzend artsen van andere disciplines, mits er regelmatig contact/samenwerking is geweest.

- *Andere werkers in de zorg*: totaal minimaal zes beoordelaars; in ieder geval beoordeling door minimaal drie verpleegkundigen. Bij voorkeur geen selectie van medische secretaresses of andere beoordelaars die weinig aanwezig zijn bij de patiëntenzorg.
- *Patiënten*: minimaal tien beoordelaars; liefst een gemengde patiëntenpopulatie van poliklinische en klinische patiënten en patiënten waarmee meerdere contacten zijn geweest (minimaal twee).

De wijze waarop het gewenste totale aantal beoordelaars per groep wordt bepaald, wordt toegelicht in hoofdstuk 4 en hoofdstuk 11C. Gezien het aantal beoordelaars dat moet worden geselecteerd heeft het de voorkeur de aios wel voldoende tijd voor de selectie te geven. Voor de betrouwbaarheid van de gegevens is een voldoende aantal beoordelaars per groep nodig en dit heeft een hogere prioriteit dan een zo kort mogelijke verwerkingstijd (hoofdstuk 10). In de Nederlandse studie is het haalbaar gebleken om in een periode van minimaal twee maanden voldoende relevante beoordelaars te kunnen selecteren, mits de setting hiervoor geschikt is (op de IC is het selecteren van voldoende patiënten bijvoorbeeld niet haalbaar). Door de MSF procedure al eerder aan te kondigen en/of op te starten, bijvoorbeeld drie tot maximaal zes maanden tevoren, kan de aios ruim vóór het beoordelingsgesprek nadenken over de vraag welke beoordelaars relevant kunnen zijn om te selecteren. Door collegae aan het begin van een klinische stage al te vragen of zij na een bepaalde periode feedback willen geven op het functioneren, krijgen zij de mogelijkheid een aios gerichter te observeren. Juist omdat het bij MSF niet gaat om korte observaties maar om een oordeel over het functioneren over een langere tijdsperiode is het belangrijk dat beoorde-

laars hierover ruim tevoren zijn geïnformeerd. Dat voorkomt uitsluitend retrospectieve en daardoor soms vervormde oordelen. De kans dat voldoende patiënten en collegae kunnen worden geselecteerd voor het geven van relevante, waardevolle feedback wordt hiermee ook vergroot.

B. Benadering van de beoordelaars

In de literatuur worden verschillende manieren beschreven waarop beoordelaars gevraagd kunnen worden door middel van vragenlijsten feedback te geven. Belangrijk is dat een zo groot mogelijke respons wordt bereikt. Voor de benadering van de beoordelaars (vragen om deelname en uitreiken van vragenlijsten) wordt soms een administratief bedrijf ingeschakeld als ‘tussenpersoon’. De artsen wordt dan gevraagd de namen, functies en adressen op te geven van een aantal geselecteerde beoordelaars. Vervolgens worden de verschillende vragenlijsten opgestuurd naar deze beoordelaars.^{3-4 6} Het medium dat gebruikt wordt heeft invloed op het ontwerp van de vragenlijsten en de snelheid waarmee data verwerkt kunnen worden. Vragenlijsten kunnen op papier ingevuld worden, via intranet, telefoon of e-mail. Iedere benaderingswijze heeft voor- en nadelen. Bij gebruik van korte vragenlijsten kan voor telefonisch invullen met cijfercodes gekozen worden, evenals bij patiënten die analfabeet zijn. Wanneer voor het invullen op de computer wordt gekozen, dienen alle beoordelaars voldoende toegang te hebben tot een computer en eventueel intranet. Bovendien zal technische ondersteuning nodig zijn voor het opzetten van digitale vragenlijsten en voor het ontwerpen van een computerapplicatie voor het verzamelen en verwerken van de data. Wanneer niet aan deze voorwaarden kan worden voldaan, heeft invullen met pen en papier of aan de hand van

een telefoongesprek de voorkeur.¹ Penny vergeleek pen en papier met elektronische systemen en concludeerde dat de methode de consistentie van beoordelaars niet beïnvloedde.³³ Archer et al. beschreven een succesvol systeem waarin beoordelingen centraal gescand worden in een computer, die vervolgens analyses uitvoert en een samenvatting maakt.³ Lockyer et al. onderzochten in 2006 de verschillen in responspercentage tussen het gebruik van telefonische vragenlijsten, invullen via internet en invullen op papier. In de eerste twee groepen waren de artsen zelf verantwoordelijk voor verspreiding van de vragenlijsten. De responspercentages op papier waren 88% voor patiënten, 94% voor andere werkers in de zorg, 91% voor medische collegae. De responspercentages voor het invullen per telefoon of internet bedroegen respectievelijk 68% en 32% voor patiënten, 62% en 38% voor andere werkers in de zorg en 57% versus 43% voor medische collegae. De verschillen in responspercentages worden deels gewijd aan het minder toegankelijk zijn van internet op het platteland. Bij het invullen aan de hand van een telefoongesprek of via internet waren de artsen verantwoordelijk voor selectie van de bronnen. Dit kan geleid hebben tot een lagere respons, omdat de werkdruk van artsen over het algemeen hoog is. De resultaten van de studie suggereren dat het invullen van vragenlijsten op papier minstens even effectief is en tot een hogere respons leidt dan technologische verspreidingsmethoden. Ondersteuning van artsen bij verspreiding van de vragenlijsten speelt volgens de auteurs een kritische rol in een succesvolle beoordelingsprocedure.²⁶ Payne waarschuwde ervoor dat men soms teveel waarde hecht aan elektronische gegevens en numerieke scores en dat het gevaar van elektronische systemen kan zijn dat subjectieve gegevens als objectief worden weergegeven.³⁴

In deze Richtlijn adviseren wij de aios zelf de beoordelaars te laten selecteren en om toestemming te vragen voor het toesturen van een vragenlijst per e-mail. Dit hoeft weinig tijd te kosten, aangezien de aios zelf samenwerkt of contact heeft met respectievelijk collegae en patiënten. Door de arts zelf de beoordelaars te laten benaderen kan op persoonlijker wijze medewerking worden gevraagd en worden de kosten van een 'tussenpersoon' bespaard. Tevens is omwille van de haalbaarheid de keuze gemaakt voor een elektronische procedure met behulp van het webbased programma *Multisourcefeedback.nl*. Dit programma is vanaf iedere computer met internetverbinding te openen met persoonlijke inloggegevens voor aios en opleiders. In het programma kunnen aios na toestemming van de beoordelaars de verkregen e-mailadressen invoeren. Beoordelaars kunnen de vragenlijst openen via een link in de e-mail die zij automatisch ontvangen nadat een aios hun e-mailadres heeft ingevoerd.

7. Anonimiteit/veiligheid van werkrelatie en arts-patiëntrelatie

Het invoeren van een procedure voor MSF kan een bedreiging betekenen voor zowel degene die beoordeeld wordt als voor de beoordelaar, maar ook de kwaliteit van de werkrelatie tussen arts en beoordelaar kan gevaar lopen. Al in het begin van de implementatie moeten daarom garanties gesteld worden om de relaties tussen collegae en artsen en patiënten te waarborgen.¹ In een vroeg stadium is dan ook de vraag gesteld hoe herkenbaar de bron mag zijn voor de aios om enerzijds te zorgen voor voldoende eerlijkheid bij het beoordelen en anderzijds geen conflicten te veroorzaken tussen de aios en collega's of de aios en patiënten. In veel studies over MSF zijn de identiteit van de beoordelaars en hun

individuele scores alleen bekend bij de opleider (of bij degene die het feedbackgesprek voert). Onder andere in de studie van Whitehouse et al. wordt op het beoordelingsformulier vereist dat de beoordelaar zijn/haar naam invult en een handtekening zet. De individuele scores worden geaggregeerd en de aios weet dus alleen dat de beoordelaar één van de beoordelaars is die hij/zij heeft gekozen.

De auteurs geven drie belangrijke redenen waarom zij vinden dat de identiteit van beoordelaars wél bekend moet zijn bij de supervisor. Ten eerste kan het voor de supervisor nodig zijn, om in gevallen waarin een beoordelaar een onvoldoende of sterk afwijkende score geeft, toelichting te vragen of verder onderzoek te doen. Dit kan alleen wanneer de opleider weet wie de beoordelaar is. Ten tweede noemen de auteurs hun ervaring met gevallen waarbij beoordelaars om persoonlijke, in plaats van professionele redenen, lage scores gaven. Die wijze van beoordelen kan ontmoedigd worden door geen volledig anonim proces te kiezen, maar te zorgen dat slechte scores herleidbaar zijn voor verder onderzoek. De derde reden die genoemd wordt voor het vragen om naam en handtekening is het voorkomen van fraude met de vragenlijsten. De auteurs geven wel aan dat verder onderzocht moet worden of collegae geremd worden in het geven van een eerlijk oordeel over artsen wanneer zij niet volledig anoniem zijn.³⁵ Er is ook een ander argument dat pleit tegen anonimiteit. Dat is de wens om een open communicatiesfeer op klinische afdelingen te creëren, mede in het kader van incidentmeldingen.

Op basis van literatuurbevindingen en discussies in het Utrechtse In VIVO kern-team wordt echter geadviseerd de individuele scores die beoordelaars hebben gegeven niet rechtstreeks bekend te maken aan de aios. De aios krijgt alleen

een gemiddelde score per groep beoordelaars te zien; een slechte score zal niet herleidbaar zijn tot een individuele beoordelaar. Geschreven opmerkingen worden per categorie beoordelaars weergegeven, zodat ook deze niet herleidbaar zijn tot een individuele persoon. Hiervoor dient het aantal beoordelaars dus minimaal twee te zijn. Geadviseerd wordt de supervisor wel inzage te geven in de individuele scores en geschreven opmerkingen. Er zal dan dus geen sprake zijn van volledige anonimiteit. In een toelichting bij de vragenlijst kan aan beoordelaars worden uitgelegd waarom dit het geval is. In hoofdstuk 13A wordt het waarborgen van de anonimiteit bij weergave van de MSF resultaten nader toegeelicht.

8. De plaats van zelfbeoordeling in de multisource feedbackprocedure

Een primair doel van MSF is de mogelijkheid voor aios om hun eigen beeld van het functioneren te vergelijken met het beeld dat anderen van hen hebben. Dit kan vooral nuttig zijn voor artsen die minder goed functioneren en mogelijk weinig zelfinzicht hebben. Daar staat tegenover dat het zelfvertrouwen van artsen, die zichzelf minder goed beoordelen, kan toenemen.⁴ Er zijn echter kanttekeningen te plaatsen bij het gebruik van zelfbeoordelingen. In veel studies worden slechte tot matige correlaties gevonden tussen zelfbeoordelingen en een beoordeling door peers, andere werkers in de zorg en door supervisors.^{10 12-13 23 26} Artsen schatten zichzelf over het algemeen lager in dan hun beoordelaars. In een Canadese studie onder artsen met een buitenlands artsdiploma schatte echter 70% van de artsen zichzelf hoger in dan hun beoordelaars.²⁶

Zelfbeoordeling of zelfreflectie in de opleiding is belangrijk, maar een zelfbeoordeling vormt in onze visie geen onder-

deel van MSF. Het zelfbeeld ontstaat juist immers mede op basis van de verkregen feedback vanuit verschillende perspectieven. Zelfbeoordeling heeft dus een andere status dan externe beoordelingen en vormt geen onderdeel van de 360 graden waarop een overall oordeel over een aios is gebaseerd. Zelfbeoordeling is in het algemeen weinig betrouwbaar en valide, maar niettemin een vaardigheid die wel ontwikkeld moet worden.³⁶ Een procedure van zelfbeoordeling kan wel heel goed als aanvulling op MSF worden toegepast. Een mogelijkheid is om de aios, voorafgaand aan het voortgangsgesprek, de vragenlijsten voor collegae en patiënten over hun functioneren ook zelf te laten invullen. De zelfbeoordeling heeft dan als doel een vergelijking te vormen met de scores die gegeven zijn door de verschillende groepen bronnen. Een grote discrepantie in de resultaten kan een gebrek aan inzicht suggereren en verschillen tussen de scores en eventuele redenen hiervoor kunnen dan besproken worden met de supervisor. Eventuele discrepanties kunnen verminderd worden door de beoordeelde aios duidelijker inzicht te laten krijgen in het gewenste gedrag. Tot slot kan de zelfbeoordeling of visie van de aios op zijn/haar competenties een belangrijke en valide toevoeging zijn aan het perspectief van de andere bronnen. Soms kunnen zelfs nieuwe aandachtspunten naar voren komen die niet door de beoordelaars zijn aangestipt.²⁵

9. Frequentie en tijdstip van multisource feedback

De frequentie van MSF is afhankelijk van de haalbaarheid en tijdsinvestering per beoordelingsronde. Om tot eventuele verbetering van het functioneren te komen moet er voldoende tijd bestaan tussen de peilingen. Frequenties van de beoordeling die worden genoemd in de relevante lite-

ratuur variëren van eenmaal per vijf jaar⁹ tot twee maal per jaar.³⁷ Geadviseerd wordt het tijdstip van een MSF procedure aan te passen aan de (vernieuwde) opleidingsschema's. Dit kan betekenen dat de procedure niet ieder jaar, maar bijvoorbeeld alleen gedurende langere klinische stages wordt doorlopen. Desgewenst, en afhankelijk van het opleidingsschema, kan de MSF ook jaarlijks worden uitgevoerd. De aios en opleider kunnen ook besluiten dit frequenter te doen; zeker bij twijfel over functioneren is dit wenselijk. In het eerste jaar van de opleiding heeft het de voorkeur om de procedure in ieder geval *éénmaal* te doorlopen.

Voor een optimale effectiviteit zou de aios deze feedback zo snel mogelijk na de beoordeling moeten ontvangen en kunnen bespreken met de opleider. De korte omlooptijd van de beoordelingen leidt tot voldoende relevantie van de feedback voor dat moment. Wanneer de beoordeling pas na lange tijd wordt meegedeeld aan de aios, zal dit voor de aios minder motiverend werken.¹ Geadviseerd wordt de periode tussen het doorlopen van de MSF procedure en het meedelen van de resultaten niet langer dan een maand te laten zijn. Idealiter staat een week na afloop van een MSF-ronde een voortgangsgesprek gepland.

10. Periode waarover beoordeeld wordt

De periode waarover een bepaalde beoordelaar de aios feedback geeft, zal per beoordelaar verschillen. Het is wenselijk dat de beoordeling van het functioneren van de aios niet slechts op basis van één enkele observatie plaatsvindt. In tegenstelling tot KPB's zal de beoordeling ten behoeve van MSF een langere periode moeten beslaan, van bijvoorbeeld minimaal één maand samenwerking bij collegae en minimaal twee arts-patiëntcontacten

bij poliklinische patiënten of bij een opname van minimaal drie dagen. Deze getallen zijn zeer arbitrair; aangezien de periode waarover beoordeeld wordt meestal niet specifiek beschreven is in de relevante literatuur. In studies over MSF in het bedrijfsleven wordt een periode van zes tot twaalf maanden op dezelfde werkplek voldoende geacht om een beoordeling te kunnen geven.³⁸ Dat lijkt lang voor een opleidings situatie, temeer omdat het advies om jaarlijks de MSF procedure uit te voeren juist is ingegeven door de gedachte dat de aios zich in een jaar substantieel ontwikkelt. Een observatieperiode van één tot drie maanden wordt daarom geadviseerd. In de meeste studies die verricht zijn in medische settings wordt ook beschreven dat de beoordelaars gedurende een langere periode moeten hebben samengewerkt met degene die zij moeten beoordelen. In de instructie voor de aios wordt benadrukt dat het belangrijk is dat de aios beoordelaars selecteert waarvan hij/zij het idee heeft dat zij relevante feedback kunnen geven.

Op de vragenlijsten voor collegae wordt gevraagd naar de periode waarover zij hun mening geven (week, maand, > 1 maand, > 3 maanden). Tijdens de verrichte studie heeft 92% van de collegae hun feedback gebaseerd op een periode van minimaal een maand of langer. Het lijkt dus haalbaar om een minimale periode van één maand als voorwaarde voor een beoordeling te stellen. De aios zal over het algemeen een aantal maanden op een afdeling werken. Wanneer er eenmaal per jaar MSF wordt gegeven zal er over het algemeen voldoende mogelijkheid zijn om beoordelaars te vinden die de aios gedurende een langere periode hebben meegemaakt. Overigens zullen er, ongeacht de periode waarover iemand de aios beoordeelt of de mate waarin de beoordelaar de aios kent, altijd beoorde-

laars zijn die niet alle vragen kunnen beoordelen, omdat zij niet alle aspecten hebben kunnen observeren.

Bij patiënten ligt dat anders. Voor de beoordeling door patiënten gaat het niet zozeer over welke periode wordt beoordeeld, maar is een voldoende aantal contacten met de arts van belang om de verschillende aspecten van het functioneren te kunnen beoordelen. Sommige vragen kunnen bijvoorbeeld alleen worden beantwoord nadat er meerdere contacten hebben plaatsgevonden tussen arts en patiënt (bijvoorbeeld deze arts komt gemaakte beloften en/of afspraken na). Bij patiënten kan daarom beter worden gevraagd naar het aantal contacten dat heeft plaatsgevonden met de aios. Het aantal contacten tussen arts en patiënt varieerde in onze pilotstudie van 1 tot >10, met een gemiddelde van 5,3 voor de Kindergeneeskunde en 3,6 voor de Gynaecologie. Dit is deels te verklaren door een groter aantal poliklinische patiënten in de Gynaecologiepopulatie. Vooral van de patiënten die slechts éénmaal contact hadden gehad met de aios, bleken een groter deel van de vragen met 'NVT' te hebben ingevuld. Een aantal maal worden opmerkingen gemaakt als: 'dit moet nog blijken', 'kan ik na één afspraak nog niet beoordelen', etc. Het lijkt wenselijk dat patiënten worden geselecteerd die minimaal twee contacten hebben gehad met de arts, echter de haalbaarheid hiervan is afhankelijk van de opleiding, setting en lengte van de klinische stages.

De periode die beoordelaars krijgen om de vragenlijsten in te vullen en te retourneren wordt in de meeste artikelen over MSF niet beschreven. Violato et al. en Davis beschrijven dat in hun studies een periode van twee weken voldoende was voor het invullen en retourneren van de vragenlijsten.^{10 23} Idealiter wordt de beoordelingsprocedure zo ontworpen dat de feedback

zo snel mogelijk na het invullen van de vragenlijsten wordt ontvangen door de beoordeelde arts. Een snelle verwerkingstijd waarborgt het nut van de feedback en stimuleert de motivatie van de beoordeelde artsen.³⁸ Het invullen van de vragenlijsten hoeft niet meer dan tien minuten te kosten en liefst worden deze nog dezelfde dag ingevuld en opgestuurd. In de digitale procedure wordt de beoordelaars gevraagd binnen een week de vragenlijst te voltooien.

Door de MSF procedure zo vroeg mogelijk aan te kondigen en/of op te starten, drie tot maximaal zes maanden tevoren, kan de aios ruim voor het beoordelingsgesprek al collegae vragen of zij over een bepaalde tijd feedback willen geven op zijn/haar functioneren. Collegae krijgen hierdoor de mogelijkheid een aios gericht te observeren en kunnen daardoor mogelijk meer relevante en waardevolle feedback geven en hoeven zich niet alleen op retrospectieve oordelen te baseren.

11. Ontwerp van geschikte beoordelingsinstrumenten

A. Uniformiteit vragenlijsten

In sommige studies wordt eenzelfde formulier beschreven voor gebruik door alle beoordelaars (peers, andere werkers in de zorg, consulenten). In verschillende (pilot) studies wordt echter beschreven dat het gebruik van verschillende vragenlijsten, aangepast voor de te ondervragen bron, een hogere betrouwbaarheid en validiteit heeft. Een nadeel van het gebruik van meerdere vragenlijsten is echter dat gegevens moeilijker geaggregeerd kunnen worden. Voor de groepen medische collegae en andere werkers in de zorg is in onze eerste pilotstudie gebruik gemaakt van dezelfde vragenlijsten, met als doel een eenvoudigere verwerking van de gegevens en analyses. Naar aanleiding van een evaluatie van de vragenlijsten door beoorde-

laars en de deelnemende aios zijn de vragenlijsten aangepast. De tweede versie van de vragenlijsten voor medische en andere werkers in de zorg bevatten 19 overeenkomstige vragen (zie hoofdstuk 11B, voor nadere toelichting).

In studies over MSF waarbij patiënten betrokken zijn in de procedure zijn specifieke vragenlijsten ontworpen voor deze categorie. Vaak komen de vragen aan patiënten wat betreft de inhoud grotendeels overeen met de vragen aan collegae. Er bestaan echter ook altijd specifieke vragen die alleen aan patiënten gesteld kunnen worden. Het lijkt de auteurs wenselijk om voor patiënten een aparte vragenlijst te gebruiken, aangezien zij de competenties communicatie en professionaliteit vanuit een ander perspectief kunnen beoordelen dan collegae. Over beide gebieden worden acht vragen gesteld aan patiënten. Het resultaat van de door patiënten ingevulde vragenlijsten zal dus een gemiddelde score leveren voor de competenties communicatie en professionaliteit, gezien vanuit het perspectief van patiënten.

Op basis van literatuuronderzoek, resultaten van de pilotstudie en overleg met experts zijn drie verschillende vragenlijsten ontworpen. Alle groepen beoordelaars wordt gevraagd naar tips of opmerkingen om het functioneren van de arts te verbeteren. De vragenlijsten zijn onderzocht op hun validiteit en betrouwbaarheid, zoals beschreven in hoofdstuk 11C, waarna zij verder zijn aangepast. Voor de inhoud van de vragenlijsten wordt verwezen naar Bijlage II van deze Richtlijn, waarin de aangepaste vragenlijsten voor iedere categorie zijn opgenomen.

B. Type vragen en beoordelingschaal

Een noodzakelijke stap in de ontwikkeling van beoordelingsinstrumenten is het bepalen van het te meten domein of gedrag.

Wanneer men besluit professionaliteit te meten, moet er bijvoorbeeld naar gestreefd worden alleen dat ‘concept’ van gedrag te meten en niet iets anders (zie paragraaf C, voor meer toelichting op inhouds- en constructvaliditeit). De vragen zelf moeten dus overeenkomen met het te meten concept of construct en het liefst gebaseerd zijn op een theoretische onderbouwing, zoals bij de in de VS, Canada en het Verenigd Koninkrijk gebruikte formulieren (zie Tabel in Bijlage I). Omdat men te maken heeft met professionals en hun oordeel over anderen, is input van de ‘eindgebruikers’ over de inhoud en formulering van de beoordelingsinstrumenten belangrijk. Deze input kan men bijvoorbeeld verkrijgen door focusbijeenkomsten of enquêtes te houden. Door het evalueren van de vragenlijsten met de verschillende groepen gebruikers kunnen de vragen steeds meer worden aangepast aan wat in de praktijk beoordeeld kan worden en aan wat de aios graag beoordeeld ziet door de verschillende groepen. Aanpassingen in bijvoorbeeld de formulering van de vragen kunnen als ontwikkelingsproces worden gezien. Door het opdoen van ervaring met de vragenlijsten zullen beoordelaars hier meer gerichte feedback op kunnen geven.¹

Vragen aan patiënten dienen logischerwijs anders geformuleerd te worden dan vragen aan collegae. Niet alleen gaan de vragen over andere aspecten van het functioneren, zij moeten ook in eenvoudigere bewoordingen worden gesteld en uitgaan van het perspectief van de patiënt. Het is uiteraard belangrijk vragen te stellen over gedrag dat door patiënten geobserveerd kan worden. De patiënt zal bijvoorbeeld geen goed oordeel kunnen vormen over de intercollegiale samenwerking of over de competentie kennis en wetenschap. Per vraag moet slechts naar één aspect van het functioneren worden gevraagd. Belangrijk is dat ook een categorie: ‘niet te beoorde-

Tabel 5. Voorbeelden van antwoordschalen.

Type	Categorieën					
Eens met stellingen	Sterk oneens	Oneens	Niet eens/ niet oneens	Mee eens	Sterk mee eens	(Onvoldoende informatief)
Frequentie		Nooit	Soms	Meestal	Altijd	(Niet van toepassing)
Verwachtingen		Voldoet niet aan verwachting	Voldoet aan verwachting	Overtreft verwachting		
Kwaliteit		Onvoldoende	Voldoende	Goed	Excellent	
Vergelijking met andere artsen	Hoort bij de slechtsten	Hoort bij onderste helft	Gemiddeld	Hoort bij bovenste helft	Hoort bij de besten	

len' of 'geen mening' wordt toegevoegd, dit om te voorkomen dat beoordelaars een ongefundeerde uitspraak doen.^{10 23} Er is geen 'juist' aantal of type antwoordmogelijkheden op een schaal te noemen, wel is informatie beschikbaar over de factoren die de selectie van een schaal kunnen beïnvloeden.³⁹

Sommige auteurs beschrijven een aanpassing van de beoordelingsschaal na evaluatie. Bij het instrument van Hesketh et al. werd een vierpuntsschaal voor kwaliteit gebruikt (excellent, good, satisfactory, requires help/attention) en was er een categorie 'niet van toepassing'. Een nadeel van deze schaal bleek dat er slechts één categorie bestond voor onvoldoende beoordelingen (alleen requires help/attention), waardoor beoordelaars minder snel geneigd zijn deze categorie aan te kruisen (lijkt op pass-fail beoordeling). Een beter alternatief dat wordt genoemd is een vijfpuntsschaal met als vijfde categorie: 'poor'.⁵ Tabel 5 toont een aantal voorbeelden van mogelijke beoordelingsschalen voor MSF.^{1 9}

Voor een gemakkelijke verwerking en vergelijking van scores adviseren wij om voor alle vragenlijsten stellingen te gebruiken, die te beantwoorden met een vijf-punts Likertschaal, bestaand uit categorieën voor de mate van overeenstemming (eens/oneens), 'niet te beoordelen' of 'niet van toepassing'. De stellingen in het programma *Multisourcefeedback.nl* zijn gebaseerd op de CanMEDS rollen die door de KNMG zijn aangepast tot zeven algemene competentiegebieden van de medische specialist,³⁰ verdeeld in vier subcompetenties. Zoals eerder beschreven is een keuze gemaakt voor een aantal competenties. Voor de stellingen zijn zoveel mogelijk de door de KNMG gebruikte bewoordingen gehanteerd. Waar nodig zijn subcompetenties omgezet in een beschrijving van specifiek gedrag.

Open vragen, bijvoorbeeld: 'wat vindt u een zeer positief aspect van het functioneren van deze arts-assistent?' geven vaak waardevolle informatie over de context waarin kwantitatieve gegevens geïnterpreteerd worden; ook hebben veel deelnemers

in eerdere studies aangegeven dat ze geschreven opmerkingen de meest zinvolle aspecten van MSF vinden. Higgins beschrijft in 2004 vragenlijsten met 45 gesloten vragen en drie open vragen: 1) Wat zijn de sterkste punten van deze arts? 2) Welke specifieke aanpassing(en) in het gedrag zullen het meeste voordeel opleveren? en 3) Welke andere opmerkingen heeft u, die de arts mogelijk kunnen helpen bij het verbeteren en/of ontwikkelen van zijn/haar functioneren?³⁷ Uit de resultaten van de Nederlandse studie kan ook geconcludeerd worden dat kwalitatieve opmerkingen voor de aios van grote waarde zijn. Iedere categorie beoordelaars wordt dan ook gevraagd naar één of meer tips of opmerkingen om het functioneren van de arts (nog meer) te verbeteren.

Het gebruik van kwantitatieve vragenlijsten bij MSF heeft als voordeel dat relatief gemakkelijk aanpassingen gedaan kunnen worden, vooral wanneer de vragenlijsten voor een formatief doeleinde gebruikt worden.¹ In veel studies wordt beschreven dat de beoordelingsinstrumenten vaak meerdere keren aangepast worden. Violato beschrijft in 2006 een evaluatie van vragenlijsten, die zijn ingevuld om kinderartsen in opleiding te beoordelen. Wanneer bij items door meer dan 20% van beoordelaars 'niet te beoordelen' was ingevuld, werden deze vragen veranderd of weggelaten in een volgende versie van de formulieren.¹¹ Ook voor deze Richtlijn is deze aanpak gebruikt voor verdere aanpassing van de vragenlijsten. In de in Nederland verrichte studie bleken drie vragen aan medische collegae en twee vragen aan andere werkers in de zorg door meer dan 20% van de respondenten te zijn beantwoord met 'niet te beoordelen'. Deze vragen zullen in een volgende versie worden aangepast. Voor geen van de vragen aan patiënten werd dit percentage gehaald.

C. Validiteit en betrouwbaarheid van de vragenlijsten

In deze paragraaf worden de begrippen inhouds- en constructvaliditeit en betrouwbaarheid toegelicht. Vervolgens worden de resultaten van onze studie op deze aspecten besproken.

In MSF procedures moet het bewijs voor validiteit opgebouwd worden om te bepalen of de beoordelingsinstrumenten het vermogen hebben te meten wat ze behoren te meten. Het onderzoeken van de validiteit van de instrumenten is een continu doorlopend proces. De totale vragenlijsten, de verschillende factoren en de subscores per competentie worden hierbij geanalyseerd. Omdat ook de score per vraag van belang kan zijn voor reflectie en voor het opstellen van een actieplan, wordt de validiteit van iedere losstaande vraag getoetst.¹ De content validity of inhoudsvaliditeit van een instrument bepaalt of de toetsmethode representatief is voor het gehele competentiegebied dat men wil meten. Deze inhoudsvaliditeit wordt in de regel gecreëerd bij de constructie van het instrument. In dit geval is het raamwerk van de CanMEDS competenties hiervoor gebruikt (zie hoofdstuk 11B). Constructvaliditeit staat voor de mate waarin het meetinstrument het construct (de competentie) meet dat men wil meten. Om te bepalen of hetzelfde construct wordt gemeten met de verschillende vragenlijsten kunnen correlaties tussen de scores van verschillende categorieën beoordelaars berekend worden. Men verwacht dan lagere correlaties tussen methodes die verschillende constructen meten, en hogere correlaties tussen methodes die verondersteld worden hetzelfde te meten. Een hoge correlatie kan er ook op wijzen dat hetzelfde wordt gemeten met twee verschillende methodes, waardoor één van de beoordelingsmethodes mogelijk achterwege kan blij-

ven. Tevens kunnen correlaties aantonen of gegevens van verschillende groepen beoordelaars juist verschillende en mogelijk aanvullende informatie opleveren. Matige correlaties ($r = .30$ to $.50$) kunnen erop duiden dat met de nieuwe manier van beoordelen andere competenties worden beoordeeld dan met bestaande methoden. Johnson en Cujec correleerden bijvoorbeeld cognitieve kennistoetsen met beoordelingen door collega-artsen en verpleegkundigen om te bepalen of deze beoordelingen overeenkwamen.²⁴ Lipner et al. correleerden hun data met eerdere beoordelingen van de algemene klinische competenties door opleiders.¹³ Zelfbeoordeling en beoordeling door peers, evenals beoordelingen door verpleegkundigen, worden ook vaak aan elkaar gecorreleerd. Correlatiestudies in medische settings hebben aangetoond dat zelf-versus peer-beoordeling lage of niet-significante correlaties oplevert ($r < .25$), terwijl collegae versus andere groepen collegae meestal matig correleren.^{8 20 24} Deze gegevens bevestigen het belang van het verkrijgen van gegevens uit verschillende groepen beoordelaars.¹ Een hoge correlatie tussen beoordelingen van de verschillende categorieën zou erop duiden dat beoordeling door slechts één van de categorieën volstaat.

Behalve het bepalen van de validiteit van vragenlijsten is het van belang de betrouwbaarheid te meten. Hoewel de betrouwbaarheid van vragenlijsten op verschillende manieren bepaald kan worden, wordt meestal in eerste instantie de interne consistentie van het gehele instrument bepaald met Cronbach's alpha. Met deze methode kan bepaald worden in hoeverre verschillende items in een meetinstrument die eenzelfde kenmerk beogen te meten, dat ook daadwerkelijk doen. De interne consistentie van het meetinstrument is groot, als alle met dit doel opge-

nomen vragen hetzelfde kenmerk meten. Voor toepassingen van MSF in medische settings is aangetoond dat met gegevens die verkregen zijn met behulp van vragenlijsten, hoge waarden voor interne consistentie kunnen worden bereikt. Met de beoordelingsinstrumenten gebruikt door het CPSA-PAR Programma zijn bijvoorbeeld Cronbach's alpha waarden van boven 0.90 bereikt.^{8-11 40}

Als tweede of alternatieve bepaling van de betrouwbaarheid van een beoordelingsinstrument wordt in veel studies de generaliseerbaarheidstheorie toegepast. Deze benadering maakt een schatting mogelijk van de betrouwbaarheid met een verschillend aantal beoordelaars en verschillende aantallen vragen door het aantal beoordelaars te variëren.^{8-11 13 40} In hoofdstuk 4 is dit type onderzoek en de resultaten van verschillende onderzoeken reeds toegelicht.

Voor de start van het implementatieproces behoren de beoordelingsinstrumenten dus op meerdere wijzen getoetst te worden. Over het algemeen geldt dat hoe zwaarder de beoordeling meeweegt, hoe belangrijker het is om voldoende bewijs te leveren voor de betrouwbaarheid en validiteit van de instrumenten.¹

D. Resultaten studie OOR Utrecht en Amsterdam

Met het eerder beschreven webbased programma is een pilotstudie verricht onder 70 aios Kindergeneeskunde en Gynaecologie/Obstetrie in de OOR Utrecht en Amsterdam (VU en UvA). Bij in totaal drie academische en acht perifere ziekenhuizen tussen juni en augustus 2008 werden totaal 1016 vragenlijsten ingevuld (zie Tabel 6). Deze studie was gericht op het bepalen van de betrouwbaarheid en validiteit van de vragenlijsten in deze Richtlijn en het bepalen van het gebruiksgemak en de haalbaarheid van de procedure. Gemid-

delde scores en responspercentages werden bepaald en de interne consistentie en intraclass correlaties werden berekend. De resultaten van de evaluatie van de haalbaarheid en het gebruiksgemak van de procedure volgen op korte termijn.

Gezien de eerste resultaten en het lage percentage gemelde problemen met het webbased programma (1%) lijkt de conclusie gerechtvaardigd dat de procedure op grote schaal haalbaar is. Er was sprake van voldoende bereidheid tot medewerking; er wordt weinig tijdsinvestering gevraagd van de betrokkenen en de informatie is rijk gebleken. Responspercentages (aantal beantwoorde uitnodigingen) bedroegen 90, 87 en 72% voor respectievelijk medische collegae, andere werkers in de zorg en patiënten. Ook het aantal gegeven tips en opmerkingen, die aios vaak zeer zinvol achten, was hoog: 62% voor patiënten tot 79% voor medische collegae. Van de 70 aios die deelnamen aan

de studie hadden echter slechts 16 aios een complete dataset van minimaal zes medische collegae en zes andere werkers in de zorg. De belangrijkste reden voor het lage gemiddelde aantal beoordelaars die door aios is gemeld, is de korte periode die aios kregen voor het verzamelen van de MSF van 22 beoordelaars. Deze periode bedroeg zes weken en vond daarnaast plaats aan het begin en tijdens de zomerperiode. Aios die op de IC (Kindergeneeskunde) werkten, net waren gewisseld van afdeling of op veel wisselende afdelingen werkten, meldten problemen met het vinden van voldoende beoordelaars. In het begin van de studieperiode bleek er ook enige weerstand bij het vragen van patiënten, enerzijds tegen het vragen van feedback op hun functioneren, anderzijds tegen het vragen om een e-mailadres. Positieve reacties van andere aios en/of het geven van nadere informatie over MSF/de procedure heeft tot een toename van het aantal

Tabel 6. Gedetailleerde resultaten respons MSF pilotstudie.

	Totaal aantal aios (%)	Totaal aantal beoordelaars (respons, %)	Gemiddeld aantal beoordelaars (range)	Medische collegae (respons, %)	Andere werkers in de zorg (respons, %)	Patiënten (respons, %)
				Gemiddelde per aios (range)	Gemiddelde per aios (range)	Gemiddelde per aios (range)
Totale respons	70 (78%)	1016 (83%)	14,5 (5-29)	420 (89,9%) 6,0 (2-11)	338 (86,9) 4,8 (1-14)	256 (71,7%) 3,7 (0-11)
Aios Kindergeneeskunde	33 (47%)	504 (84%)	15,3 (8-29)	190 (88,8%) 5,8 (2-10)	165 (89,2%) 5,0 (1-14)	147 (72,8%) 4,5 (0-11)
Aios Gynaecologie/Obstetrie	37 (53%)	512 (84%)	13,8 (5-25)	230 (90,9%) 6,2 (3-11)	173 (84,8%) 4,7 (0-7)	109 (70,3%) 3,0 (0-10)

uitnodigingen geleid in het tweede deel van de studieperiode. Voor de verzamelperiode wordt naar aanleiding van deze resultaten drie maanden geadviseerd.

Over de beoordelingen van deze collegae op overeenkomstige vragen over één aios werden intraclass correlaties (ICC) berekend voor de scores van alle collegae samen en voor de twee groepen apart. In Tabel 7 worden deze ICC's weergegeven in een correlatiematrix. Deze tabel bevat ICC's van beoordelingen van 16 aios, gegeven door totaal minimaal 12 collegae op overeenkomstige vragen voor medische en andere werkers in de zorg. Over de beoordelingen door patiënten zijn geen ICC's berekend aangezien er geen overeenkomstige vragen aan collegae werden gesteld.

In deze voorlopige versie van de Richtlijn zijn, tot een nadere aanvulling volgt, de resultaten van de eerste pilot weergegeven in een matrix (Tabel 7). In deze matrix zijn de correlaties die als hoog zouden moeten gelden volgens bovenstaande hypothese gearceerd. Uit de matrix is op te maken dat de correlaties van dezelfde CanMEDS constructen vanuit de perspectieven van medici, niet-medici en patiënten niet consequent hoger zijn dan correlaties van verschillende constructen binnen de categorieën of over de categorieën van beoordelaars. Bovendien zijn de correlaties tussen de categorieën nergens hoog, met uitzondering van 'organisatie'. Met andere woorden: de perspectieven van medici, niet-medici in de zorg en van patiënten op dezelfde competentiegebieden zijn klaarblijkelijk verschillend, zoals gemeten met deze vragenlijsten, met uitzondering van het competentiegebied organisatie. Deze resultaten komen overeen met eerdere verrichte studies.¹⁻²

Zoals in Tabel 8 is weergegeven is de interne consistentie (Cronbach's alpha)

van afzonderlijke competentiegebieden voor beoordelaarsgroepen, voor zover deze berekenbaar was, voldoende voor uitspraken per competentiegebied, c.q. per rapportage van scores per competentiegebied.

12. Logistieke verwerking

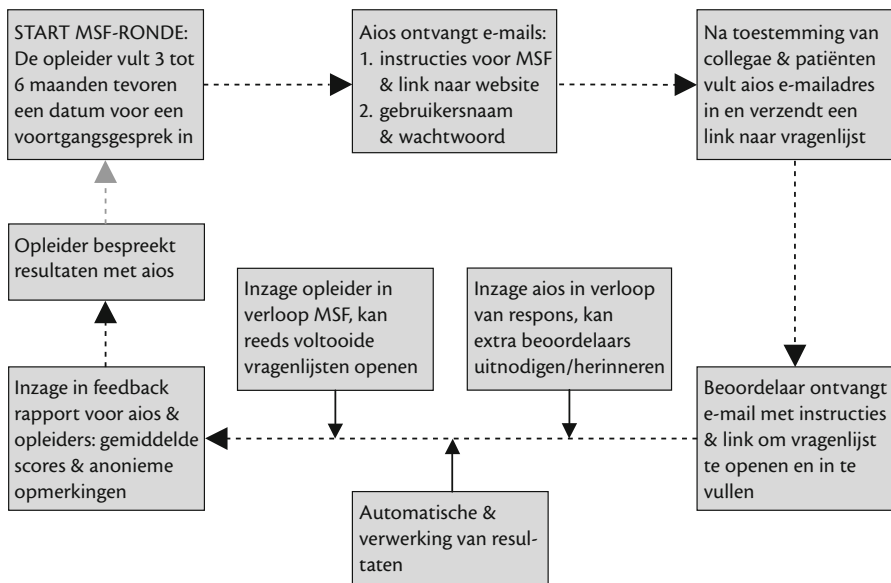
Uit de relevante literatuur blijkt dat voor een voldoende haalbaarheid van de procedure de tijdsinvestering van alle betrokkenen niet te hoog moet zijn. Zoals beschreven in hoofdstuk 6, de benadering van bronnen, maken medische organisaties in het buitenland veelal gebruik van elektronische programma's dan wel van een extern onderzoeksbureau voor de uitvoering van MSF procedures. De MSF procedure, zoals voorgesteld in deze richtlijn, is op enige schaal slechts uitvoerbaar met een elektronisch programma dat administratie en gegevensverwerking voor zijn rekening neemt. Het programma *Multisourcefeedback.nl* is met steun van het CBOG tot stand gekomen en is vervaardigd om toegepast te kunnen worden in alle medische vervolopleidingen in Nederland. Met dit webbased programma kunnen:

- opleiders de feedbackprocedure voor aios in gang zetten en monitoren;
- aios uitnodigingen versturen naar beoordelaars;
- beoordelaars vragenlijsten invullen;
- geanonimiseerde feedbackrapporten worden geraadpleegd door aios en opleiders;
- opleiders ingevulde vragenlijsten volledig inzien.

In Figuur 1 is de procedure zoals deze verloopt met behulp van *Multisourcefeedback.nl* in een stroomdiagram weergegeven.

Tabel 8. Interne consistentie vragenlijsten MSF pilotstudie.

Vragenlijst/categorie	Competentiegebied	Cronbach's α
Medische collegae	Medisch handelen	0.72
	Communicatie	0.73
	Samenwerking	0.80
	Organisatie	0.72
	Kennis en wetenschap	0.64
	Professionaliteit	0.85
Andere werkers in de zorg	Communicatie	0.85
	Samenwerking	0.84
	Organisatie	0.79
	Professionaliteit	0.86



Figuur 1. Stroomdiagram.

A. Logistieke taken voor (plaatsvervangende) opleiders

Zowel academische als perifere (plaatsvervangende) opleiders kunnen een account aanvragen bij de beheerder van het programma met vermelding van naam, instelling, specialisme en te gebruiken e-mailadres. De (plaatsvervangende) oplei-

der moet geregistreerd staan bij de beroepsvereniging. Met de per e-mail ontvangen persoonlijke inloggegevens kan worden ingelogd op de homepage van het programma, waar een handleiding te vinden is voor (plaatsvervangende) opleiders. Aios moeten door de opleider eenmalig worden geregistreerd met hun geboorte-

datum, opleidingsjaar en e-mailadres. Eventueel kan deze administratieve taak worden overgedragen aan een secretaresse. Wanneer een aios wordt toegevoegd worden automatisch inloggegevens verstuurd naar de aios en is zijn/haar account geactiveerd. Afhankelijk van de werkplek kan een MSF ronde worden gestart door de perifere of de academische opleider. Dit dient maximaal zes maanden en minimaal twee maanden vóór het gesprek te gebeuren, zodat de aios genoeg tijd heeft om voldoende beoordelaars te selecteren. Nadat de procedure is gestart ontvangt de aios automatisch een e-mail met instructie voor het uitvoeren van de procedure en een link naar de site om beoordelaars uit te nodigen. In Box 2 is de inhoud van deze e-mail weergegeven.

B. Logistieke taken voor aios

Aios kunnen hun opleider vragen naar de mogelijkheid van het gebruik van *Multi-sourcefeedback.nl*. Zodra aios zijn toegevoegd in het programma, ontvangen zij persoonlijke inloggegevens. Na het starten van een MSF ronde door de opleider ontvangen aios bovenstaande e-mail zodat zij beoordelaars kunnen uitnodigen hen feedback te geven. Dit kan door het e-mailadres van collegae en patiënten in te voeren, nadat de aios om hun toestemming voor medewerking hebben gevraagd (zie later). Wanneer er geen e-mailadres is kan de aios zo nodig het eigen e-mailadres invullen, of dat van een afdelingssecretaresse. De direct ontvangen e-mail bevat naast een link naar de vragenlijst ook inloggegevens. Door de e-mail uit te printen en bijvoorbeeld aan een patiënt mee te geven kan diegene met deze gegevens vanaf een willekeurige computer met internetverbinding de vragenlijst openen via de website. Tijdens een MSF ronde kunnen aios in hun persoonlijke account zien wie al geantwoord heeft (niet hun

Box 2. Instructie voor uitvoering procedure.

Geachte mevrouw/de heer aios,

Volgens de opgave die is ontvangen heeft u op dd-mm-jjjj a.s. een voortgangsgesprek. Ter voorbereiding hierop dient een Multi-Source Feedback procedure te worden uitgevoerd.

Het is hiervoor noodzakelijk dat u een aantal personen uit uw omgeving vraagt om u met behulp van een vragenlijst feedback te geven op uw functioneren:

- a. zes tot acht **medische collegae** (minimaal één stafflid en één collega aios)
- b. zes tot acht **andere werkers in de zorg** (minimaal twee verpleegkundigen)
- c. tien tot twaalf **patiënten** van de polikliniek of afdeling die u minimaal twee keer gezien heeft

U wordt verzocht om uiterlijk één week voor het voortgangsgesprek alle e-mailadressen op te geven van de betreffende collegae en patiënten. U kunt dit doen door op de onderstaande link te klikken. De beoordelaars ontvangen van ons een korte vragenlijst. U kunt op de website www.multisourcefeedback.nl zien welke beoordelaars u hebt uitgenodigd en wanneer een vragenlijst is voltooid. Ook kunt u beoordelaars een herinnering versturen of de beoordelaar verwijderen voordat de vragenlijst is voltooid. Uiterlijk vier dagen voor het voortgangsgesprek is op de website voor u en uw opleider een overzicht te zien van de gemiddelde scores en gegeven tips onder 'Voltooid feedbackprocedures'. MSF link: Klink hier om collegae en patiënten uit te nodigen.

Met vriendelijke groeten, Beheerder MSF

antwoorden) en zo nodig nieuwe beoordelaars uitnodigen of herinneringen versturen. Op de homepage zijn instructies voor de aios en een toelichting voor beoordelaars te vinden. Het samenvattende feedbackrapport dat na afloop van een MSF ronde beschikbaar is op de website, kan geprint worden voor een handmatig portfolio.

C. Afronden van de feedbackprocedure

De procedure wordt automatisch één week vóór de datum van het voortgangsgesprek afgerond. Dan wordt een feedbackrapport opgemaakt, ter inzage voor zowel de aios als de opleider. Bij eventuele (logistieke) problemen tijdens de 'verzamelpe-

riode' kan verlenging van deze periode nodig blijken; aan de opleider kan dan gevraagd worden de datum naar voren te schuiven. Verlenging is dus alleen mogelijk wanneer de datum in *Multisourcefeedback.nl* wordt gewijzigd door de opleider.

D. Inzage in vertrouwelijke gegevens en bewaartermijn

Het programma is een beveiligde website waartoe alleen toegang wordt verkregen met inloggegevens die beheerd worden door de programmabeheerder. Zowel opleiders als aios hebben een persoonlijke toegang. Alleen de opleider die de MSF ronde heeft gestart, heeft inzage in de afzonderlijke vragenlijsten die tijdens een MSF ronde zijn ingevuld. Academische opleiders hebben inzage in de gegevens en feedbackrapporten van de aios uit zijn/haar OOR. Perifere opleiders hebben alleen inzage in de feedbackrapporten van aios die op dat moment werkzaam zijn in hun instelling.

13. Van multisource feedback naar concrete feedback

A. Weergave van de multisource feedback

De resultaten van de beoordelingen kunnen op verschillende manieren worden weergegeven.

In veel bestaande MSF procedures worden de scores van verschillende beoordelaars geaggregeerd, ofwel samengevat weergegeven voor de beoordeelde arts (soms per groep bronnen; soms is er van alle beoordelaars één gemiddelde score). De over het algemeen anonieme verwerking en weergave geldt zeker voor de door patiënten ingevulde vragenlijsten, aangezien de arts-patiëntrelatie niet beïnvloed dient te worden. Het verdient de voorkeur ook de beoordelingen door collegae anoniem te verwerken (zie hoofdstuk 7). Hoevel geaggregeerde scores de anonimiteit

waarborgen, is het nadeel dat belangrijke verschillen in perspectieven moeilijker te herkennen zijn. Idealiter worden de scores van verschillende groepen beoordelaars apart weergegeven, waarbij zowel het gemiddelde als de spreiding van scores worden getoond, zodat verschillen in perspectieven toch enigszins zichtbaar zijn. Wanneer voldoende beoordelaars worden gebruikt, zal de anonimiteit gehandhaafd blijven. Vrije tekstboodschappen moeten zorgvuldig worden weergegeven (niet herleidbaar tot de beoordelaar) en in sommige omstandigheden zelfs misschien gecensureerd.¹ In sommige MSF procedures worden de scores van een individuele arts weergegeven samen met de gemiddelde scores van andere beoordeelde artsen.^{3 6 9-11}

De wijze waarop de arts de resultaten ontvangt en kan bespreken, verdient ook enige aandacht. Begeleiding rondom het ontvangen van de feedback is essentieel en het verdient dan ook de voorkeur om de resultaten niet te lang voor het geplande beoordelingsgesprek te laten inzien. De feedback komt van collegae en opleidingsgenoten met wie de aios nog zal samenwerken in de toekomst. In de zeldzame gevallen dat er onverwacht een probleem in het functioneren blijkt te zijn, moet het voor de aios mogelijk zijn dit direct met de opleider te bespreken. In de praktijk leidt de MSF procedure vrijwel altijd tot een positieve beoordeling door vele collegae en patiënten.¹

Door het digitale programma *Multisourcefeedback.nl* wordt na afronding van de MSF, minimaal één week voorafgaand aan het (jaarlijkse) voortgangsgesprek, een samenvattend rapport opgemaakt voor aios en opleider. In het 'feedbackrapport' worden de gemiddelde scores per categorie beoordelaars (respectievelijk medische collegae, andere werkers in de zorg en patiënten) en het aantal res-

pondenten per categorie getoond. De scores per groep en competentiegebied worden ook in staafdiagrammen weergegeven. Door op een score voor een competentiegebied te klikken, worden de gemiddelde scores per afzonderlijke vraag weergegeven. De tips voor verbetering per categorie beoordelaars worden geanonimiseerd weergegeven; hiervoor zijn minstens twee gegeven tips per groep vereist. Mogelijk kan er op termijn voor worden gekozen de gemiddelde resultaten van de aios te vergelijken met de scores van een referentiegroep, bestaande uit aios van hetzelfde opleidingsniveau. In Bijlage III is een voorbeeld te zien van een feedbackrapport in het programma *Multisourcefeedback.nl*. De verkregen resultaten van zo'n MSF ronde vormen als het ware 'het eigendom' van de aios, maar blijven in te zien door de opleider. De rol die beiden spelen in een MSF procedure wordt toegelicht in de volgende twee paragrafen.

B. Rol van de aios

In de voorgestelde MSF procedure is de aios verantwoordelijk voor het selecteren van de juiste bronnen, het vragen van medewerking aan beoordelaars en voor het toesturen van de vragenlijsten aan beoordelaars. Aangezien het functioneren van de aios beoordeeld wordt, heeft het de voorkeur hem/haar verantwoordelijk te laten zijn voor het verkrijgen van voldoende ingevulde vragenlijsten per groep beoordelaars (totaal minimaal 22 vragenlijsten) om een betrouwbare beoordeling te kunnen krijgen.

Voor het verkrijgen van waardevolle feedback is het in het belang van de aios om zo relevant mogelijke beoordelaars te vragen, ofwel collegae waarmee geruime tijd is samengewerkt en patiënten waarmee zoveel mogelijk contact is geweest.

Aios krijgen uiterlijk een week voordat de resultaten worden besproken in een

voortgangsgesprek inzage in het feedbackrapport. Tijdens dit gesprek krijgt de aios de kans om de opleider toelichting te vragen op de resultaten. Van de aios wordt een actieve rol verwacht in de voorbereiding op en tijdens het gesprek. Wanneer de beoordeling de persoonlijke ontwikkeling tot doel heeft, is het opstellen van specifieke leerdoelen een belangrijk onderdeel in de feedbackmethode.¹ Eén van de deelnemende opleiders aan de Nederlandse studie heeft positieve ervaringen met het stellen van de volgende twee vragen aan de aios ter voorbereiding op het voortgangsgesprek:

- Welke punten van de MSF wil je meenemen wat betreft feedback op je *gedrag* en hoe wil je aan die punten gaan werken?
- Welke punten van de MSF wil je meenemen wat betreft feedback op je *persoon* en in hoeverre denk je hier verandering in aan te (moeten en/of kunnen) brengen?

Aios kunnen door hun opleider gevraagd worden zichzelf te beoordelen aan de hand van de items op de vragenlijst voor zowel patiënten als collegae. Tijdens het voortgangsgesprek kunnen de resultaten van deze zelfbeoordeling vergeleken worden met de andere resultaten. Wanneer geen gebruik wordt gemaakt van een zelfbeoordeling kan het gesprek worden voorbereid door bijvoorbeeld te vragen 'wat gaat goed volgens de drie groepen', 'wat kan beter' en 'hoe kan ik deze punten verbeteren'. De leerdoelen kunnen volgens de SMART criteria opgesteld worden (specifiek, meetbaar, acceptabel/haalbaar, resultaatgericht en tijdgebonden).

Whitehouse et al. beschrijven in hun studie over MSF dat de opleider na het voortgangsgesprek een samenvatting maakt van de MSF en een eindoordeel geeft dat gebaseerd is op de resultaten van de vragenlijst-

ten, op eventueel nader onderzoek (bijvoorbeeld toelichting door beoordelaars) en op bespreking van de resultaten met aios.³⁵ Het advies is om deze samenvatting niet door de opleider, maar door de aios te laten maken. De aios kan dan direct een vergelijking maken van de sterke en zwakke punten met zijn/haar zelfbeoordeling. Deze analyse en de onderliggende gemiddelde beoordelingen en geschreven opmerkingen worden opgenomen in het portfolio.

C. Rol van de opleider

MSF kent fundamentele verschillen met andere vormen van feedback. De geschikte persoon om de feedback aan de beoordeelde arts te verstrekken varieert afhankelijk van het doel van de MSF. In de MSF procedure zoals voorgesteld in deze Richtlijn, wordt geadviseerd de opleider de rol van feedbackgever te geven. Zowel de aios als de opleider krijgen voorafgaand aan een beoordelingsgesprek, of in ieder geval na afronding van een feedbackprocedure, een samenvattend rapport van de verkregen feedback. De opleider zal geen rol hebben in de selectie van beoordelaars, maar heeft wel zicht op de beoordelaars die door de aios gevraagd zijn een MSF vragenlijst in te vullen. Er wordt geadviseerd de opleider enkele weken voor het voortgangsgesprek te laten checken of de aios initiatief onderneemt en de tevoren gemaakte afspraken nakomt. De verwerking en samenvatting van de resultaten verloopt automatisch, zodat opleiders meer tijd kunnen besteden aan het eventueel vragen van toelichting aan beoordelaars en het op de juiste wijze bespreken van de resultaten met de aios. De opleider schift de eventuele extreem alarmerende boodschappen uit het samenvattende rapport. Deze kunnen eventueel met de betreffende collega worden doorgesproken. Het resultaat van

deze terugkoppeling kan eventueel met de beoordeelde arts worden besproken; anonimiteit van de bron blijft hierbij echter vereist. De opleider kan door deze rol dus bepalen welke weging er wordt gegeven aan de verzamelde gegevens. Uit zowel de gemiddelde scores als uit de geschreven opmerkingen kan een aantal sterke en zwakke punten worden gedestilleerd (mogelijk ook gebleken uit andere beoordelingsmethoden). Vervolgens wordt een grote lijn van feedback bepaald en met de aios besproken in een persoonlijk gesprek. De feedback dient zoveel mogelijk met expliciete voorbeelden te worden geïllustreerd.

Het nut van het gebruik van MSF is sterk afhankelijk van de vaardigheden van de opleider om feedback te geven. Idealerweise ontvangt hij/zij dan ook een training of instructie voor het geven van feedback volgend op de resultaten van de MSF.³ Het geven van feedback moet onderscheiden worden van het geven van een beoordeling en het doel van MSF moet duidelijk zijn voor de opleider. Er wordt anoniem feedback gegeven op gedrag en de feedback kan daarmee worden gezien als de beoordeling van iemand als persoon, zowel door degene die feedback geeft als door degene die ontvangt. Door een gebrek aan voorbereiding van één of beide partijen kan het bespreken van de feedback leiden tot misverstanden of erger. De verkregen feedback kan berichten opleveren die erg moeilijk zijn om te ontvangen (bijvoorbeeld collega-aios die het nemen van verantwoordelijkheid voor eigen handelen slecht beoordelen). De feedbackgevers, in dit geval de opleiders, moeten goed geïnformeerd zijn over het beleid van de organisatie ten aanzien van MSF en het doel en de werkwijze van de procedure. Ook moeten zij gewend raken aan de vragenlijsten en daaruit voortvloeiende gegevens. Verder moet men bewust wor-

den van de verschillende reacties die mensen kunnen vertonen op het ontvangen van feedback en de vaardigheden die nodig zijn voor het houden van een constructief feedbackgesprek. Een vaardige feedbackgever zal een individu kunnen helpen om conclusies te trekken en deze te koppelen aan verschillende personen en situaties.⁴¹

In de literatuur worden belangrijke aspecten genoemd van het geven van feedback naar aanleiding van MSF. Vier hoofdprincipes voor het geven van feedback, ook toepasbaar bij MSF, worden door Holmboe et al. beschreven in een review over feedback naar aanleiding van korte klinische beoordelingen, die ook hier nuttig kunnen zijn. Het feedbackgesprek dient gericht te zijn op het vaststellen van sterke punten en op het helpen van de aios te bepalen op welke punten verbetering nodig is.⁴² Vermeden moet worden om zich tijdens het gesprek te veel te richten op geïsoleerde negatieve opmerkingen of uitzonderingen. Beoordelaars zouden constructieve in plaats van destructieve feedback moeten geven, maar desondanks kunnen de geschreven opmerkingen iemand doen schrikken of overstuur maken. Adequate ondersteuning en verdere begeleiding zijn in deze situaties zeer belangrijk. Verder is het belangrijk te onthouden dat de MSF suggereert welke competenties verbeterd kunnen worden, maar dat nader onderzoek nodig is om de precieze aard en ernst van de problemen te bepalen. Bij een lage gemiddelde score voor communicatie met patiënten kunnen bijvoorbeeld andere beoordelingsstechnieken, zoals video-opname, gebruikt worden om de precieze aard van het probleem vast te stellen.⁴ Tot slot is het belangrijk dat wanneer probleemgebieden naar voren zijn gekomen er door de opleider en aios een duidelijk actieplan wordt gemaakt door leerdoelen vast te leggen in een per-

soonlijk ontwikkelingsplan en dat dit plan navolging krijgt in vervolgesprekken.⁴²

De resultaten van de MSF moeten in combinatie met andere beoordelingsmethoden worden geïnterpreteerd. Desondanks biedt deze manier van beoordelen op zichzelf voor artsen in opleiding een waardevolle kans om gestructureerde feedback te krijgen van uiteenlopende collega's. Op die manier kunnen zij meer inzicht krijgen in hun eigen rol en functioneren als professional, wat een zinvolle professionele ontwikkeling vergemakkelijkt.⁴ Wanneer opleiders meerdere feedbackgesprekken hebben gehouden, kunnen zij waardevolle informatie verschaffen over de thema's die veel voorkwamen tijdens de gesprekken. Hierbij moet natuurlijk wel de privacy van individuen gehandhaafd blijven. Deze thema's kunnen een stimulans vormen om breder gedragen problemen in de opleiding aan te pakken.⁴¹

14. Opname gegevens in het portfolio

Gedurende de opleiding worden de samenvattende rapporten van voltooide feedbackprocedures bewaard in het programma *Multisourcefeedback.nl*. Het heeft de voorkeur een papieren of digitale versie van het samenvattend rapport van een MSF ronde op te nemen in het portfolio. Na het voortgangsgesprek zijn de voltooide vragenlijsten niet meer afzonderlijk in te zien, tenzij er naar aanleiding van het voortgangsgesprek behoefte blijkt te zijn aan nader onderzoek. Om zelfreflectie te stimuleren kan de sterke/ zwakte analyse die de aios maakt naar aanleiding van o.a. de multisource feedback ook een verplicht onderdeel van het portfolio vormen. Tot slot wordt geadviseerd een sterk negatief ingevulde vragenlijst op te nemen in het dossier dat de opleider van de aios heeft.

Het MSF programma is niet ontwikkeld als onderdeel van een specifiek elektro-

nisch portfolio. Niettemin leent het zich goed voor koppeling aan een dergelijk systeem. Het Maastricht-Manchester Toets-service e-portfolio model in ontwikkeling biedt thans de mogelijkheid om rapportages van *Mutisourcefeedback.nl* automatisch te integreren in de overall rapportage voor aios.

15. Beschreven effecten van multisource feedback op het functioneren van artsen

Studies tonen aan dat veel artsen de gegevens van MSF willen gebruiken om veranderingen in hun handelen en functioneren te maken, maar de mate waarin de gegevens gebruikt worden kan worden beïnvloed door tal van interne en externe factoren. De interne factoren omvatten de eigen perceptie van artsen van hun prestaties, emoties, persoonlijke verwachtingen, en van hun persoonlijke waarden en capaciteit ten aanzien van verandering. De externe factoren die de arts beïnvloeden omvatten de aard, geloofwaardigheid, specificiteit en de consistentie van de feedback vergeleken met andere feedback en de barrières voor verandering.⁴³ In dit kader is het belangrijk om realistische verwachtingen te hebben over verandering, aangezien er aanzienlijke variatie bestaat in hoe informatief de gegevens zijn voor de persoon, evenals in de waargenomen noodzaak tot verandering.⁴⁴

Er zijn tot op heden nog geen longitudinale studies gepubliceerd in de medische setting, die het effect op scores gedurende twee of meer tijdspannes onderzochten. Dit is echter wel onderzocht in het bedrijfsleven en in een meta-analyse van 24 longitudinale studies waarbij meer dan twee metingen verricht werden, en waar aangetoond werd dat de meetbare verbetering over het algemeen klein was, maar meer waarschijnlijk was voor sommige 'ontvangers' dan voor anderen.⁴⁵ Verbetering was

meer waarschijnlijk wanneer gegevens een noodzaak tot verandering suggereerden, ontvangers een positieve visie hadden op feedback, zelf de behoefte hadden om hun gedrag te veranderen, positief reageerden op de feedback, geloofden in haalbaarheid van verandering, toepasselijke doelstellingen maakten om hun gedrag aan te passen en wanneer actie werd ondernomen tot verbetering van vaardigheden en prestaties. Dit soort analyses van het gebruik van de gegevens door de feedbackontvangers zijn belangrijk bij het proberen te verkrijgen van voldoende bewijs voor validiteit van de beoordelingsinstrumenten. Wanneer mensen hun feedbackgegevens negeren, is het wel erg moeilijk om voldoende argumenten voor geldigheid van de procedure te vergaren.

Brinkman et al. rapporteerden in 2007 de resultaten van een gerandomiseerde gecontroleerde trial waarin het effect van MSF op communicatieve vaardigheden en professionaliteit van kinderartsen in opleiding werd onderzocht.⁴⁶ Achttien aios kindergeneeskunde werden tweemaal volgens een MSF procedure beoordeeld door ouders van patiënten en door verpleegkundigen; zij vulden een zelfbeoordeling in en kregen een coachingsessie, naast de bestaande standaard feedback. In de controlegroep kregen 18 aios alleen de standaardfeedback; tot het einde van de trial werden zij geblindeerd voor de resultaten van beoordelingen door verpleegkundigen en ouders. Beide groepen werden tweemaal beoordeeld, met een tussenliggende periode van vijf maanden. De beoordelingen door patiënten verbeterden in beide groepen; in de eerste groep leken de scores meer verbeterd, er werden echter geen statistisch significante verschillen gevonden. De beoordelingen door verpleegkundigen bleven in de eerste groep gelijk of verbeterden, maar de scores in de tweede groep

verslechterden. De verschillen in scores waren statistisch significant en suggereerden een positief effect van MSF op het gebied van effectief communiceren met patiënten en familie, het op tijd vervullen van taken en het demonstreren van verantwoordelijkheid en aansprakelijkheid. Deze studie heeft dus aangetoond dat de interventie MSF een positieve invloed kan hebben op de competenties communicatie en professionaliteit.

16. Implementatie van de multisource feedback en mogelijke weerstanden

Ten aanzien van het daadwerkelijk invoeren van MSF is in de voorgaande hoofdstukken al een aantal adviezen gegeven. In dit hoofdstuk worden belangrijke aspecten van de implementatie van MSF en de eventuele weerstanden bij invoering van het proces besproken.

Tot op heden bestond er in de medische vervolgopleidingen geen gestructureerde beoordeling van het functioneren van aios door collega's waarmee nauw wordt samengewerkt of door patiënten. MSF zal dus voor vrijwel alle betrokkenen onbekend zijn en een goede voorlichting over de procedure is dan ook van belang. Wat betreft schriftelijke voorlichting is behalve deze Richtlijn ook een folder beschikbaar over het programma *Multisourcefeedback.nl*, en te downloaden van de betreffende website. Tijdens trainingen voor opleiders, stafleden en aios kan MSF mondeling toegelicht worden en kan worden geoefend met het bespreken van de resultaten door de opleider met de aios. Betrokken beoordelaars worden bij het digitaal invullen van de vragenlijsten zo duidelijk mogelijk voorgelicht over het doel van de MSF, over de anonieme weergave van resultaten en over de beveiliging van gegevens. Het lijkt logistiek niet haalbaar om alle beoordelaars te trainen.

Behalve goede voorlichting zijn twee andere belangrijke praktische uitdagingen in de ontwikkeling van een MSF procedure – het ontwerpen van geschikte beoordelingsinstrumenten voor alle beoordelaars en het organiseren van de gegevensverzameling van een grotere groep individuen – om tot een geaggregeerd feedbackrapport te komen, waarbij vertrouwelijkheid van gegevens wordt gewaarborgd.⁴⁷ In hoofdstuk 7 is de anonimiteit van beoordelaars besproken. De keuzes die hierbij gemaakt zijn hebben als doel de weerstand te verkleinen ten aanzien van de mogelijk negatieve invloed van MSF op persoonlijke werk- of arts-patiëntrelaties. Met een automatische procedure kan MSF op grote schaal worden uitgevoerd en kan een te hoge tijdsinvestering van betrokkenen worden voorkomen. Ten behoeve van een logistieke verwerking met een lage tijdsinvestering voor betrokkenen is een digitaal programma ontworpen, waarmee een studie is verricht onder 70 aios. De geschiktheid van de beoordelingsinstrumenten is onderzocht aan de hand van de grote aantallen vragenlijsten die verkregen werden in deze studie. De vragenlijsten zijn voldoende valide en betrouwbaar gebleken voor een formatief doel.

Het proces van implementatie van deze nieuwe procedure in de bestaande culturen en organisaties vergt tijd. Voor de invoering en aanpassing van de procedure kan een periode van twee tot drie jaar nodig zijn. Goede MSF procedures vereisen steun en toewijding van de leiders van organisaties of afdelingen, goede communicatie met de deelnemers aan de procedure over doel, werkwijze en gegevensverwerking en tot slot een gemotiveerd aansturend kernteam dat het implementatieproces kan begeleiden.¹ Van steun en vertrouwen van opleiders, stafleden, aios en overige betrokkenen ten aanzien van

MSF als formatieve feedbackmethode zal in Nederland in sommige gevallen al sprake zijn, aansturing lijkt dan minder nodig. In andere gevallen zal deze nog moeten groeien en sommige betrokkenen zullen om verschillende redenen de methode permanent wantrouwen. Opleiders kunnen aios stimuleren door zichzelf ook open te stellen voor feedback, bijvoorbeeld om feedback te vragen op zijn/haar opleiderskwaliteiten. Tijdens de implementatie of eerste ronde kan tegen verschillende weerstanden worden aangekomen bij alle groepen betrokkenen. Aios kunnen bijvoorbeeld moeite hebben met de acceptatie van het beoordeeld worden door collega's van hun eigen of andere disciplines. Zij kunnen bang zijn voor slechte beoordelingen van collega's/patiënten (waarmee zij misschien ook nog een nauwe band hebben) en voor de eventuele maatregelen bij een slechte beoordeling. Ook kunnen de beoordeelde artsen bang zijn voor het publiek worden van deze gegevens, of voor een onvoldoende vertrouwelijke behandeling van gegevens.² Een gebrek aan vertrouwen in de anonieme verwerking van gegevens kan bij beoordelaars leiden tot een te hoge beoordeling of weerstand in het geven van tips ter verbetering.¹ Beoordelaars kunnen moeite hebben met het geven van een eerlijk en oprecht oordeel over het functioneren van (collega-) artsen. Wanneer deelnemers meer vertrouwd raken met het hele proces, neemt de angst en weerstand ten aanzien van de procedure vaak af en wordt het vermogen om de procedure effectief toe te passen meestal groter.³⁷ Bij de ontwikkeling en invoering van MSF is het tevens belangrijk zich te realiseren dat de feedback slechts een onderdeel is van meerdere beoordelingsprocessen gedurende de opleiding. In tegenstelling tot reeds gestandaardiseerde beoordelingsmethoden kunnen kleine variaties bij de invoer-

ring van het proces echter tot grote verschillen leiden in de kwaliteit van de informatie die wordt verkregen. Het is dan ook van belang rekening te houden met mogelijke weerstanden van deelnemers en deze bespreekbaar te maken.

In de toepassing van *Multisourcefeedback.nl* procedure bij de opleiding Gynaecologie/Obstetrie in de Onderwijs- en opleidingsregio Amsterdam/AMC is met succes één maand per jaar uitgeroepen tot 'MSF-maand'; in de weken daarvoor doorlopen alle aios de procedure. In het andere half jaar wordt een vergelijkbare feedbackprocedure uitgevoerd waarin de aios de klinische stafleden beoordelen. Van deze complementaire procedure gaat een grote kracht uit van een open en kwetsbare opstelling waarvan zowel aios als de klinische staf beter worden.

Bij het werken met MSF zijn de visies van de 'eindgebruikers' en van degenen die worden beoordeeld een essentieel onderdeel van de implementatie en evaluatie. Verschillende studies tonen aan dat zowel de beoordelaars als beoordeelde artsen MSF zien als een belangrijk beoordelingsinstrument.^{5 9 22 29 32} Higgins et al. beschrijven in 2004 dat participerende aios positiever reageren op MSF naarmate de procedure langer bestaat. Volgens de auteurs kan dit secundair zijn aan de angst die er in eerste instantie bestond om door uiteenlopende collegae te worden beoordeeld. Ook geven zij aan dat iedere verandering in de cultuur van een afdeling of organisatie leidt tot weerstanden, die echter de neiging heeft met het vorderen van de tijd te verminderen. Dit is zeker het geval wanneer een positieve cyclus van verandering in gang gezet wordt, gekenmerkt door verbetering van de omgang tussen aios en andere werknemers in de zorg.³⁷ In de acceptatie van het proces kan meespelen dat de feedback die wordt verkregen met MSF, in tegen-

stelling tot toetsing van kennis en vaardigheden, zeer gevoelig van aard kan zijn. Wanneer de feedback op het functioneren niet goed wordt gecommuniceerd of weergegeven, kan dit direct of op lange termijn ernstige gevolgen hebben. Artsen werken vaak gedurende langere tijd samen met verschillende collegae. Het ontdekken dat collegae hun prestaties als onvoldoende of suboptimaal beoordelen, kan een grote verrassing zijn en tot onbedoelde reacties leiden, zelfs tot lang nadat de feedback is gegeven^{1 32 44} Onvoldoende aandacht voor de beoordelaars, de setting waarin de procedure plaatsvindt en voor de wijze waarop feedback wordt gegeven kan de bruikbaarheid van MSF ernstig verminderen.¹

17. Multisource feedbackprocedures voor de minder patiëntgebonden specialismen

Verreweg het meeste onderzoek over MSF voor medische opleidingen is uitgevoerd voor patiëntgebonden specialismen. Voor deze specialismen is deze Richtlijn en het programma *Multisourcefeedback.nl* primair bedoeld. Uiteraard kan in de opleiding in andere medische specialismen, waaronder radiologie, medische microbiologie en anaesthesiologie MSF evenzeer nuttig zijn. *Multisourcefeedback.nl* kan ook daar worden toegepast, maar dan met openlating van de categorie van de patiënten, zodat twee groepen van beoordelaars overblijven.

Literatuur

- Lockyer J, Clyman SM, Holmboe E, Hawkins R (eds) in: A Practical Guide to the Evaluation of Clinical Competence, Elsevier, 2007 (MSF Chapter), in press.
- Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof* 2003 Winter; 23(1):4-12.
- Archer J, Norcini J, Davies HA. Peer review of paediatricians in training using SPRAT. *BMJ* 2005; 330:1251-53.
- Davies HA, Archer JC. Multi source feedback using Sheffield Peer Review Assessment Tool (SPRAT): Development and practical aspects. *Clinical Teacher* 2005; 2(2):77-81.
- Hesketh EA, Anderson F, Bagnall GM, Driver CP, Johnston DA, Marshall D, et al. Using a 360 degrees diagnostic screening tool to provide an evidence trail of junior doctor performance throughout their first postgraduate year. *Med Teacher* 2005; 27(3):219-33.
- Archer J, Norcini J, Southgate L, Heard S, Davies H. Mini-PAT (peer assessment tool): A valid component of a national assessment programme in the UK? *Adv in Health Sciences Education* 2006 [DOI 10.1007/s10459-006-9033-3].
- Woolliscroft JO, Howell JD, Patel BP, Swanson DB. Resident-patient interactions: The humanistic qualities of internal medicine residents assessed by patients, attending physicians, program supervisors and nurses. *Acad Med* 1994; 69(3):216-224.
- Violato C, Marini A, Toews J, Lockyer J, Fidler H. Using peers, consulting physicians, patients, co-workers and self to assess physicians. *Acad Med* 1994;72 (10 S):57-63.
- Hall W, Violato C, Lewkonja R, Lockyer J, Fidler H, Toews J., et al. Assessment of physician performance in Alberta: The Physician Achievement Review Project. *CMAJ* 1999;161 (1):52-57.
- Violato, C, Lockyer, JM, Fidler H: Multi source feedback. A method of assessing surgical practice, *BMJ* 2003;326: 546-48.
- Violato C, Lockyer J, Fidler H. The assessment of pediatricians by a regulatory authority, *Paediatrics* 2006;117:79-802.
- Violato C, Marini A, Toews J, Lockyer J, Fidler H. Feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. *Acad Med* 1997; 72(10 Suppl 1):S82-4.
- Lipner RS, Blank LL, Leas BF, Fortna GS: The value of patient and peer ratings in recertification, *Acad Med* 2002;77 (10S):64-66.
- College of Physicians and Surgeons of Alberta, Physician Achievement Review Program: www.par-program.org
- Evans R, Elwyn G, Edwards A. Review of instruments for peer assessment of physicians. *BMJ* 2004; 328:1240-45.
- Mazor K, Clauser BE, Holtman M, Margolis MJ. Evaluation of Missing Data in an Assessment of Professional Behaviors. *Acad Med.* 2007;82(10 Suppl):S44-S47.
- Epstein RM. Assessment in medical education. *N Engl J Med.* 2007 Jan 25;356(4):387-96.
- Crossley J, Eiser C, Davies HA. Children and their parents assessing the doctor-patient interaction: a rating system for doctors' communication skills. *Med Educ* 2005;39(8):820-8.

19. Norcini JJ. Peer assessment of competence: *Med Educ* 2003; 37(6):539-43.
20. Ramsey PG, Carline JD, Inui TS, Larson EB, LoGerfo JP, Wenrich MD. Use of peer ratings to evaluate physician performance. *JAMA* 1993; 269:1655-60.
21. Rezler AG, Bruce NC, Schmitt BP, et al. Dilemmas in the evaluation of residents. Proceedings of the Annual Conference on Resident Medical Education 1986; 25:371-378.
22. Wenrich MD, Carline ID, Giles LM, Ramsey PG. Ratings of the performances of practicing internists by hospital-based registered nurses. *Acad Med* 1993; 68:680-87.
23. Davis JD. Comparison of faculty, peer, self, and nurse assessment of obstetrics and gynecology residents. *Obstet Gynecol.* 2002 Apr; 99(4):647-51.
24. Johnson D, Cujec B. Comparison of self, nurse and physician assessment of residents rotating through an intensive care unit. *Critical Care Medicine* 1998; 26:1811-1816.
25. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. *Med Teach* 2006 Nov;28(7):185-91.
26. Lockyer J, Blackmore D, Fidler H, Crutcher R, Salte B, Shaw K, Ward B, Wolfish N. A study of a multi-source feedback system for international medical graduates holding defined licences. *Med Educ* 2006 Apr; 40(4):340-7.
27. Butterfield PS, Mazzaferri EL. A new rating form for use by nurses in assessing residents' humanistic behavior: *J Gen Intern Med* 1991; 6:155-61.
28. Ramsey PG, Wenrich MD. Use of professional associate ratings to assess the performance of practicing physicians: past, present, future. *Adv Health Sci Educ Theory Pract* 1999; 4(1):27-38.
29. Wood L, Wall D, Bullock A, Hassell A, Whitehouse A, Campbell I. 'Team observation': a six-year study of the development and use of multi-source feedback (360-degree assessment) in obstetrics and gynaecology training in the UK. *Med Teach* 2006 Nov; 28(7):177-184.
30. Bleker OP, Cate Th.J. ten, Holdrinet RSG. De algemene competenties van de medisch specialist in de toekomst. *Tijdschrift voor Medisch Onderwijs* 2004;23(1):4-14.
31. Ten Cate TJ. Trust, competence, and the supervisor's role in postgraduate training. *BMJ* 2006; 333:748-751
32. Sargeant JM, Mann KV, Ferrier SN, Langille DB, Muirhead PD, Sinclair DE. Responses of rural family physicians and their colleagues and coworker raters to a multi-source feedback process: a pilot study. *Acad Med* 2003;78 (S10):42-44.
33. Penny JA. Exploring differential item functioning in a 360-degree assessment: rater source and method of delivery. *Organisational Research Methods* 2003;6(1):61-79.
34. Payne, T. Editorial: 360 degree assessment and feedback. *International Journal of Selection and Assessment* 1998; 6(1): 16-18.
35. Whitehouse A., Hassell A., Wood L., Wall D., Walzman M., Campbell I. Development and reliability testing of TAB – a form for 360° assessment of senior house officers' professional behaviour, as specified by the General Medical Council. *Med Teach* 2005;27: 252-58.
36. Eva K, Regehr G. Self-Assessment in the Health Professions: A Reformulation and Research Agenda. *Acad Med* 2005;80(10):S46-S54
37. Higgins RS, Bridges J, Burke JM, O'Donnell MA, Cohen NM, Wilkes SB. Implementing the ACGME general competencies in a cardiothoracic surgery residency program using 360-degree feedback. *Ann Thorac Surg* 2004 Jan;77(1):12-17.
38. Bracken DW, Timmreck CW, Church AH (Eds): *The Handbook of Multisource Feedback: The Comprehensive Resource for Designing and Implementing MSF Processes*. San Francisco: Jossey-Bass, 2001.
39. Cox EP. The optimal number of response alternatives on a scale. *J Marketing Research* 1980; 17:407-22.
40. Lockyer JM, Violato C. An examination of the appropriateness of using a common peer assessment instrument to assess physician skills across specialties. *Acad Med.* 2004 Oct;79(10 Suppl): S5-8.
41. British Psychological Society: [www.psychtesting.org.uk/\[zoek op: 360 degree feedback\]](http://www.psychtesting.org.uk/[zoek op: 360 degree feedback]).
42. Holmboe ES., Yepes M, Williams F, Huot SJ, Feedback and the mini clinical evaluation exercise. *J Gen Intern Med*, 2004. 19(5 Pt 2):558-61
43. Fidler H, Lockyer J, Toews J, Violato C. Changing physicians' practices: The effect of individual feedback. *Acad Med* 1999;74:702-14.
44. Sargeant J: *Multi-source feedback for physicians learning and change*. Maastricht University. Doctoral dissertation. September 2006.
45. Smither JW, London M, Reilly RR: Does performance improve following multisource feedback? A theoretical model, meta-analysis and review of empirical findings. *Personnel Psychology* 2005; 58:33-66.
46. Brinkman WB, Geraghty SR, Lanphear BP, Khoury JC, Gonzalez del Rey JA, Dewitt TG, Britto MT. Effect of multisource feedback on resident communication skills and professionalism: a randomized controlled trial. *Arch Pediatr Adolesc Med* 2007 Jan;161(1):44-9.
47. Toolbox: Accreditation Council for Graduate Medical Education Competencies: www.acgme.org/outcome/comp/compFull.asp

De auteurs:

Prof. dr. Th.J. ten Cate is hoogleraar medische onderwijskunde en directeur Expertisecentrum voor Onderwijs en Opleiding, UMC Utrecht.

Drs. M.A. Horsman is huisarts in opleiding en was te tijde van dit project werkzaam bij het Expertisecentrum voor Onderwijs en Opleiding, UMC Utrecht.

Correspondentieadres:

Prof. dr. Th.J. ten Cate, Expertisecentrum voor Onderwijs en Opleiding, UMC Utrecht, postbus 85500, 3508 GA Utrecht. E-mail: t.j.tencate@umcutrecht.nl

Belangenconflict: geen gemeld

Financiële ondersteuning: geen gemeld