



**Erkenningscommissie
Gedragsinterventies Justitie
*Ministerie van Veiligheid en Justitie***

**criterium 10
van de beoordelingscriteria
effectiviteit gedragsinterventies**

Advies
11 januari 2011

Colofon

Afzendgegevens **Erkenningscommissie Gedragsinterventies Justitie**

Kalvermarkt 53
2511 CB Den Haag
Postbus 20301
2500 EH Den Haag
www.erkenningscommissie.nl

Contactpersoon

W.M. Kleiman
senior beleidsmedewerker

T 070 370 61 41
erkenningscommissie@minjus.nl

Inhoud

Samenvatting - 7

- 1 Inleiding - 9**
- 2 Effectiviteitsonderzoek en het huidige criterium 10 - 11**
- 3 Ervaringen met criterium 10 - 15**
 - 3.1 Ervaringen van indieners en onderzoekers - 15
 - 3.2 Ervaring Erkeningscommissie beoordeling criterium 10 - 16
 - 3.3 Overige kanttekeningen - 17
- 4 Voorstel wijziging criterium 10 - 21**
 - 4.1 Nieuw criterium 10 - 21
 - 4.2 Onderbouwing - 21
- 5 Conclusie - 25**

Samenvatting

Naar aanleiding van de procesevaluatie van de Erkenningscommissie Gedragsinterventies Justitie, heeft de toenmalige Staatssecretaris van Justitie in 2009 aangegeven in overleg te zullen gaan over een aantal aandachtspunten, waaronder de volgende:

- of in de huidige beoordelingscriteria de aansluiting op de uitvoeringspraktijk in Nederland in voldoende mate is meegenomen en of de criteria op dit punt aanpassing behoeven;
- of met de huidige beoordelingscriteria voldoende rekening kan worden gehouden met gedragsinterventies voor kleine doelgroepen.

Het Ministerie van Justitie / DGP JS heeft de Erkenningscommissie verzocht om in het kader van deze twee aandachtspunten, te adviseren over criterium 10.

Voor de totstandkoming van onderhavig advies is een analyse gemaakt van de scores op de beoordelingscriteria en zijn gesprekken gevoerd met indieners, onderzoekers en ontwikkelaars. Ook is criterium 10 herhaaldelijk onderwerp van bespreking geweest in de Erkenningscommissie. Daarnaast is de richting van het advies besproken met het WODC en het Nederlands Jeugdinstituut.

Met betrekking tot bovengenoemde aandachtspunten is de Erkenningscommissie van oordeel dat het huidige criterium 10 inderdaad te weinig aansluit op de uitvoeringspraktijk in Nederland. Ook houdt het huidige criterium onvoldoende rekening met kleine doelgroepen.

De commissie adviseert om de inhoud van criterium 10 zodanig te wijzigen dat minimaal een structureel systeem van meten en hermeten (ROM) wordt verlangd en zo mogelijk een onderzoeksdesign voor effectmeting met een experimentele groep en controlegroep. Voor kleine doelgroepen is (onderbouwd) een ander onderzoeksdesign mogelijk.

1 Inleiding

Aanleiding advies

In 2009 heeft in opdracht van het Ministerie van Justitie / DG PJS een procesevaluatie van de Erkenningscommissie Gedragsinterventies Justitie (verder: Erkenningscommissie of commissie) plaatsgevonden. De bevindingen van deze evaluatie zijn op 18 december 2010 door de toenmalige Staatssecretaris aan de Tweede Kamer aangeboden¹. Daarbij kondigde de Staatssecretaris aan met de Erkenningscommissie in overleg te zullen treden over een aantal aandachtspunten, waaronder de volgende:

- of in de huidige beoordelingscriteria de aansluiting op de uitvoeringspraktijk in Nederland in voldoende mate is meegenomen en of de criteria op dit punt aanpassing behoeven;
- of met de huidige beoordelingscriteria voldoende rekening kan worden gehouden met gedragsinterventies voor kleine doelgroepen.

Het voorliggende advies 'Criterion 10 van de beoordelingscriteria effectiviteit gedragsinterventies' kan worden beschouwd als bijdrage aan het bovenstaande overleg, waarbij gesteld moet worden dat de bovenstaande twee aandachtspunten alle beoordelingscriteria betreffen en dus verder strekken dan alleen criterium 10.

Uit de procesevaluatie komt naar voren dat criterium 10, zoals momenteel vormgegeven, voor de nodige hoofdbrekens zorgt bij ontwikkelaars, indieners en onderzoekers. Vandaar dat de Erkenningscommissie heeft gemeend bij voorrang aandacht te moeten besteden aan dit criterium. Bij brief van 19 april 2010² heeft het Ministerie van Justitie / DG PJS hiermee ingestemd. De uitvoerbaarheid van de overige criteria zal later ter hand worden genomen.

Vorbereiding advies

Ter voorbereiding van onderhavig advies is gebruik gemaakt van gegevens die verkregen zijn door middel van gesprekken met indieners en onderzoekers. Deze vonden plaats in het kader van de voorbereiding van het advies 'Vijfjaarstermijn erkende interventies en beoordeling effectiviteit'³. In de gesprekken werd, behalve aan de vijfjaarstermijn, ook aandacht besteed aan criterium 10 en de uitvoering van (effectiviteits)onderzoek. De twee onderwerpen hangen immers nauw samen.

Ook is de richting van onderhavig advies besproken met het WODC en het Nederlands Jeugdinstituut.

Tevens is er door de Erkenningscommissie op 4 november 2010 een werkbijeenkomst georganiseerd waarvoor indieners, onderzoekers, ontwikkelaars en beleidsmedewerkers zijn uitgenodigd. Tijdens deze bijeenkomst was één workshop expliciet gewijd aan criterium 10. In deze workshop vond tussen de aanwezigen, onder wie leden van de Erkenningscommissie, een uitwisseling plaats van ervaringen en ideeën over criterium 10 en (effectiviteits)onderzoek.

¹ TK 2009-2010, 28684, nr. 259

² Kenmerk DJJ/5650521/10

³ Erkenningscommissie Gedragsinterventies Justitie, 23 december 2010

Daarnaast is criterium 10 herhaaldelijk onderwerp van bespreking geweest in de Erkenningscommissie. De ervaringen met criterium 10 en de verschillende overwegingen bij een mogelijke wijziging van de inhoud daarvan, zijn zorgvuldig bediscussieerd.

Leeswijzer

Allereerst wordt in paragraaf 2 het huidige criterium 10 toegelicht. Paragraaf 3 geeft een schets van de ervaringen die zijn opgedaan met het toepassen van criterium 10 en plaatst enkele kanttekeningen bij dit criterium. Paragraaf 4 bevat een voorstel voor een gewijzigde inhoud van criterium 10. Tevens geeft het een onderbouwing voor de voorgestelde veranderingen. Tenslotte wordt in paragraaf 5 de belangrijkste conclusies weergegeven.

2 Effectiviteitsonderzoek en het huidige criterium 10

Gedragsinterventies die in het kader van een strafrechtelijke voorwaarde, straf of maatregel worden aangeboden moeten (op termijn) bewezen effectief zijn. Dat wil zeggen dat deze interventies aantoonbaar moeten bijdragen aan het terugdringen van recidive van de deelnemers.

Om de effectiviteit van deze gedragsinterventies te kunnen vaststellen is empirisch onderzoek noodzakelijk. Dit onderzoek zal duidelijk maken in hoeverre de beoogde behandeldoelen van de interventie zijn behaald en, conditioneel daarop, in hoeverre de gewenste gedragsmodificatie, in casu recidivevermindering, is opgetreden. Immers, als niet vaststaat dat de behandeldoelen worden behaald, kan eventuele recidivevermindering lastiger aan de veronderstelde werkzame bestanddelen van de gedragsinterventie worden toegeschreven.

Om de causaliteit tussen de werkbare bestanddelen van de gedragsinterventie en de opgetreden gedragsverandering te kunnen vaststellen, zal het effectiviteitsonderzoek een experimenteel karakter moeten hebben. Dat wil zeggen dat een experimentele groep vergeleken wordt met een controlegroep.

Daarnaast dient het onderzoek te voorzien in voldoende grote groepen deelnemers (justitiabelen) om uitspraken te kunnen doen over het opgetreden effect.

Ook moeten er – behalve voor en tijdens de interventie – geruime tijd na beëindiging van de gedragsinterventie metingen plaatsvinden teneinde de duurzaamheid van de opgetreden gedragsverandering te kunnen vaststellen.

Criterium 10 schrijft voor dat een gedragsinterventie vergezeld moet gaan van een plan om de effectiviteit van de gedragsinterventie te onderzoeken. Dit plan moet, zoals geformuleerd in criterium 10, in ieder geval de volgende elementen bevatten:

- Er is een onderzoeksvraag met betrekking tot het behandeldoel(en) van de interventie;
- Er is een onderzoeksvraag met betrekking tot recidive, met een uitsplitsing naar aard van het delict (gerelateerd aan de interventie) en ernst van het delict;
- Er is een onderzoeksdesign en een onderbouwing van de keuze daarvoor;
- Binnen dit design is er een beschrijving van hoe de selectie van onderzoeksgroepen zal plaatsvinden en een onderbouwing in hoeverre de grootte van de groepen een precieze effectmeting garandeert, gegeven de verwachte base rate van criminogene factoren en recidive;
- Er is een beschrijving van de te hanteren meetinstrumenten en een onderbouwde indicatie van de validiteit en betrouwbaarheid daarvan, om zowel de recidive als de verandering op de behandeldoelen gedurende de uitvoering van de gedragsinterventie in beeld te brengen;
- Er is een beschrijving hoe eventueel randomisatie wordt gegarandeerd, dan wel een onderbouwing van het oordeel van de indieners dat de groepen vergelijkbaar zijn, en een beschrijving van de mogelijkheden voor statistische controle achteraf;
- De verschillende meetmomenten worden aangegeven: er is een voormeting (gericht op de te beïnvloeden criminogene factoren), een

eerste meting direct na beëindiging van de gedragsinterventie (gericht op de behandeldoelen, c.q. te beïnvloeden criminogene factoren) en ten minste één follow-up meting na minstens twee jaar (gericht op de recidive), maar bij voorkeur meerdere follow-up metingen.

In het huidige criterium 10 is tevens opgenomen dat een procesevaluatie vooraf dient te gaan aan het effectiviteitsonderzoek.

Criterium 10 schrijft dus voor dat het onderzoeksplan onder meer voorziet in een vergelijkbare controlegroep. Daarbij wordt gerefereerd aan de Maryland Scientific Methods Scale (MSMS)⁴. Op deze schaal wordt de kwaliteit van de onderzoeksopzet van een effectevaluatie en daarmee de bewijskracht van het onderzoek geclassificeerd aan de hand van vijf niveaus. Hoe hoger het niveau, des te groter de bewijskracht van het onderzoek. Tabel 1 geeft de vijf niveaus weer.

Tabel 1 Maryland Scientific Methods Scale

1. studies waarin gekeken wordt naar de samenhang tussen de aanwezigheid van een programma en de uitkomstvariabele na het volgen van het programma (= meting op één bepaald punt in de tijd);
2. studies waarin per deelnemer vóór en ná het programma metingen worden gedaan van het aantal gepleegde strafbare feiten en eventueel ook van de soorten door hem gepleegde vergrijpen (= veranderingsonderzoek);
3. studies met een voor- en nameting waarbij het delinquentieniveau van de onderzoeksgroep wordt vergeleken met dat van een vergelijkbare groep die het programma niet heeft gevolgd, de controlegroep (= quasi-experimenteel onderzoek);
4. studies met een voor- en nameting waarbij het delinquentieniveau van de onderzoeksgroep wordt vergeleken met dat van een controlegroep en waarbij sprake is van statistische controle voor de invloed van eventuele verschillen tussen beide groepen (= quasi-experimenteel onderzoek);
5. studies met een voor- en nameting waarbij het delinquentieniveau van de onderzoeksgroep wordt vergeleken met dat van een controlegroep en waarbij de kandidaten volstrekt willekeurig (gerandomiseerde controlegroep⁵) aan één van beide groepen zijn toegewezen (= experimenteel onderzoek).

De huidige normering met betrekking tot criterium 10 is zodanig dat er 0 punt wordt gescoord als er geen controlegroep is; dus als het een onderzoeksplan betreft op MSMS niveau 1 of 2 (of minder). Erkenning of Voorlopige erkenning van de betreffende gedragsinterventie is daardoor niet mogelijk⁶. Als het onderzoeksplan voorziet in een controlegroep op

⁴ Sherman et al, 1997. 'Preventing Crime: What Works, What Doesn't, What's Promising'. Report to the US Congress. Washington DC: U.S. Dept. Of Justice.

⁵ Ook genoemd: Randomized Controlled Trial (RCT).

⁶ Per criterium kan een gedragsinterventie een score van 0, 1 of 2 punten krijgen. Naarmate de gedragsinterventie meer voldoet aan het betreffende criterium wordt een hogere score gegeven.

MSMS niveau 3, dan wordt er 1 punt gescoord. Is er sprake van een onderzoeksdesign van MSMS niveau 4 of 5, dan worden 2 punten behaald.

Tot zover de inhoud van het huidige criterium 10. Binnen, maar ook buiten de Erkenningscommissie⁷, is de afgelopen periode uitgebreid stil gestaan bij de inhoud van criterium 10. De ervaring met het beschrijven van dit criterium ten behoeve van een ex ante beoordeling door de Erkenningscommissie, en het uitvoeren van het betreffende voorgenomen onderzoek, zijn factoren die hierbij van belang zijn geweest. De volgende paragraaf gaat daar nader op in.

Een gedragsinterventie krijgt het oordeel Erkend als deze op alle tien criteria tezamen in totaal 16 of meer punten heeft behaald en er op geen der criteria 0 punt wordt gescoord. Een gedragsinterventie krijgt het oordeel Voorlopig erkend als deze op alle criteria tezamen 12 of meer punten heeft behaald en er op geen der criteria 0 punt wordt gescoord.

⁷ Bijvoorbeeld tijdens de werkbijeenkomst die de Erkenningscommissie op 4 november 2010 heeft georganiseerd.

3 Ervaringen met criterium 10

Zowel indieners, onderzoekers als de Erkenningscommissie hebben, ieder vanuit hun eigen rol, in de afgelopen jaren ervaring opgedaan met criterium 10. Het blijkt dat er verschillende kanttekeningen kunnen worden geplaatst bij het huidige criterium.

Deze paragraaf gaat in op een aantal ondervindingen.

3.1 Ervaringen van indieners en onderzoekers

- Gedragsinterventies worden vaak in combinatie met andere interventies gegeven. De beoogde gedragsverandering, namelijk vermindering van recidive, is veelal de resultante van verschillende gedragsinterventies die parallel of serieel aan de deelnemers worden aangeboden. Daarnaast spelen zogenaamde non-specifieke factoren een rol, zoals aspecten in de relatie tussen de trainers of behandelaars en de deelnemer(s) en het behandelklimaat in de klinische setting. Het blijkt een complexe aangelegenheid om de toegevoegde waarde van een specifieke interventie uit de beoogde gedragsverandering van het totale aanbod te isoleren.
Kern van het hier benoemde probleem is dat het erg moeilijk is om tegelijkertijd de interventie breed uit te zetten in de praktijk – implementatie in het veld – en te voldoen aan de rigoureuze condities die wetenschappelijk onderzoek stelt, waaronder het aanbieden van een pure interventie zonder ‘vervuiling’ door andere.
- Indieners geven aan dat de min of meer gelijktijdige landelijke invoering van de verschillende erkende gedragsinterventies met soms geheel of gedeeltelijk dezelfde behandelmethoden en gericht op geheel of dezelfde doelgroepen, het lastig maakt om controlegroepen te vinden. De complexiteit betreft vooral de noodzaak om met het oog op het bepalen van de effectiviteit een controlegroep te vinden die geen interventie krijgt of een interventie die minder intensief is.
- Voor interventies die zijn ontwikkeld voor kleine doelgroepen is een voldoende grote instroom voor een recidivemeting, met een experimentele en een controle conditie, moeilijk te realiseren en alleen op lange termijn haalbaar. De termijn van vijf jaar waarop gegevens over de gebleken effectiviteit van de gedragsinterventie beschikbaar moeten zijn, is voor deze interventies (mogelijk) te kort.
- Het wordt als problematisch ervaren om gerandomiseerd gecontroleerd onderzoek uit te voeren bij interventies die door de rechter worden opgelegd.
- Een belangrijk punt is de recidivemeting bij deelnemers die in een JJI, PI of kliniek verblijven. Criterium 10 schrijft voor dat er na minstens twee jaar tenminste één follow-up meting is gericht op recidive. Een recidivemeting veronderstelt dat wordt nagegaan of de deelnemer na afloop van een gedragsinterventie zich al dan niet wederom schuldig maakt aan delicten. Een gedragsinterventie die in een gesloten residentieel kader wordt gegeven, is vaak (veel) eerder afgerond dan

het gehele verblijf. Doordat het verblijf van deze justitiabelen langer duurt dan de betreffende gedragsinterventie, zijn zij na twee jaar - t.t.v. een follow-up meting - óf nog steeds in de inrichting, óf zijn nog maar kort in de vrije samenleving.

- Een onderzoeksdesign dat ten behoeve van de ex ante beoordeling van de gedragsinterventie is geschreven, kan geheel of gedeeltelijk gedateerd zijn op het moment van uitvoering van het effectonderzoek, zo leert de ervaring bij indieners en onderzoekers. Veranderde omstandigheden (bijvoorbeeld dalende instroom, afwezigheid van bepaalde gegevens) moeten soms leiden tot het maken van andere keuzen ten aanzien van het inrichten van het onderzoek. Hierdoor ontstaat het risico dat het schrijven van een onderzoeksdesign alleen dient t.b.v. het verkrijgen van de ex ante erkenning en hiermee tot een rituele dans verwordt.

3.2 Ervaring Erkeningscommissie beoordeling criterium 10

Waar het gaat om het indienen van een onderzoeksdesign conform de vereisten van criterium 10, worden de ervaringen uit de praktijk in grote lijnen weerspiegeld in de scores die voor criterium 10 worden behaald bij de beoordeling van de gedragsinterventies door de Erkeningscommissie. In totaal heeft de commissie 34 verschillende gedragsinterventies één of meerdere keren beoordeeld⁸. Als gekeken wordt naar de laatste beoordeling van deze gedragsinterventies, dan kregen deze gemiddeld genomen 1,03 punten op criterium 10. Dit is minder dan het aantal punten dat gemiddeld werd behaald op de overige negen criteria, t.w. 1,41 punten. Alleen op criterium 4 - Effectiviteit behandelmethoden - wordt aanmerkelijk lager gescoord. Dit is te verklaren uit het feit dat er op dit moment geen Nederlandse effectstudies beschikbaar zijn waarmee de effectiviteit kan worden aangetoond. Zie ook tabel 2 waarin de gemiddelde scores op de 10 criteria onder elkaar zijn gezet.

Tabel 2 Gemiddelde scores op de tien criteria bij de laatste beoordeling van 34 beoordeelde gedragsinterventies

| Criteria | Gemiddelde scores |
|--|-------------------|
| 1. Theoretische onderbouwing | 1,30 |
| 2. Selectie van justitiabelen | 1,15 |
| 3. Dynamische criminogene factoren en protectieve factoren | 1,64 |
| 4. Effectiviteit behandelmethoden | 0,82 |
| 5. Vaardigheden | 1,67 |
| 6. Fasering duur en intensiteit | 1,21 |
| 7. Motivatie | 1,64 |

⁸ De Erkeningscommissie heeft tot en met 1 juli 2010 69 keer een gedragsinterventie beoordeeld. Het komt regelmatig voor dat een gedragsinterventie meerdere keren voor beoordeling bij de Erkeningscommissie wordt ingediend.

| | |
|--------------------------|------|
| 8. Continuïteit | 1,55 |
| 9. Programma-integriteit | 1,76 |
| 10. Evaluatie | 1,03 |

Van de 34 beoordeelde gedragsinterventies is bij de laatste beoordeling 6 keer 0 punt op criterium 10 gescoord, terwijl 21 keer 1 punt en 7 keer 2 punten werd gescoord.

Van de 34 beoordeelde gedragsinterventies hebben 17 gedragsinterventies het oordeel Erkend gekregen. Van deze 17 erkende gedragsinterventies behaalden 7 gedragsinterventies 2 punten en behaalden 10 gedragsinterventies 1 punt op criterium 10. Hoewel dit in theorie mogelijk zou zijn, is het echter nog niet voorgekomen dat een score van 0 punt op criterium 10 ertoe heeft geleid dat de interventie als gevolg daarvan niet kon worden erkend.

De ervaring leert dat criterium 10 weliswaar in verhouding lager scoort, maar dat in de meeste gevallen toch één of meer punten op dit criterium worden gehaald. Het blijkt dus wel mogelijk om een goed onderzoeksdesign aan te leveren. Zoals uit de gesignaleerde knelpunten in par. 3.1 naar voren komt, ligt het probleem (mede) in de aansluiting op de praktijk van het effectiviteitsonderzoek (de bruikbaarheid / toepasbaarheid).

3.3

Overige kanttekeningen

Naast de hierboven genoemde punten, is er een aantal andere kanttekeningen te plaatsen bij het huidige criterium 10.

- Een kanttekening met een principiële karakter heeft betrekking op de legitimiteit van criterium 10 als onderdeel van het toetsingskader. Het toetsingskader bestaat uit tien criteria. Deze criteria dragen elk in meer of mindere mate bij aan de mogelijke effectiviteit van de gedragsinterventie. Zo ziet criterium 1 op de vraag of de gedragsinterventie is gegrondvest op een theoretisch kader dat algemeen wetenschappelijk aanvaard is. Criterium 2 ziet op de wijze waarop gekomen wordt tot de inclusie van die personen waarvoor de interventie bestemd is. Volgens het derde criterium moet de interventie zich richten op de veranderbare dynamische criminogene factoren en protectieve factoren. Het gebruik van effectieve behandelmethoden, de fasering, duur en intensiteit, motivatie, continuïteit en een integere uitvoering zijn stuk voor stuk zaken die bepalend zijn voor de mogelijke effectiviteit van een gedragsinterventie. Uitzondering hierop is criterium 10. Dit criterium ziet niet op de mogelijke effectiviteit van de gedragsinterventie zelf, maar op de wijze waarop de effectiviteit van de interventie het best kan worden onderzocht. De eisen die onder dit criterium vallen zijn belangrijk om uiteindelijk de effectiviteit van de interventie te kunnen vaststellen, maar zijn strikt genomen niet relevant voor het ex ante vaststellen van de effectiviteit van de interventie. Zo bezien is criterium 10 een

vreemde eend in de bijt⁹. Het is - in deze lijn van redeneren - logischer dat de eisen ten aanzien van de manier waarop het effectiviteitsonderzoek is of wordt uitgevoerd pas aan de orde komen op het moment dat de effectiviteit moet worden aangetoond; dus bij de ex post beoordeling. Dan is immers de vraag actueel welke gegevens zijn aangeleverd en op welke wijze (dus: met behulp van welk type onderzoek) deze gegevens zijn verkregen.

- Tot dit moment zijn de interventies die aan de Erkenningscommissie zijn voorgelegd, nieuw ontwikkeld en beschreven, dan wel vertalingen en bewerkingen van (effectieve) interventies uit het buitenland. Ook al is het tot nog toe nog niet het geval geweest, toch is het voorstelbaar dat zich situaties voordoen waarin bestaande interventies waarvan de effectiviteit is aangetoond, aan de commissie worden voorgelegd om te worden erkend. Een hieraan verwante mogelijkheid is, dat een andere partij dan de indiener het onderzoek wil uitvoeren of al aan het uitvoeren is. Het is met andere woorden de vraag of er altijd een effectiviteitsonderzoek moet worden geëist, en zo ja of de eisen aan het uit te voeren onderzoek in alle gevallen gelijk moeten zijn. Denkbaar is dat in het geval er nog niets bekend is over de gebleken effectiviteit van een gedragsinterventie hogere eisen aan het uit te voeren onderzoek moeten worden gesteld dan in gevallen waarbij de effectiviteit inmiddels gebleken is. Dit zou een flexibele invulling van criterium 10 verlangen.
- Het huidige criterium 10 stelt enkel de effectiviteit van de gedragsinterventie vast voor een omschreven onderzoeksgroep. Onderzoek zou niet beperkt moeten blijven tot die ene onderzoeksgroep. Voor het zicht krijgen op blijvende effectiviteit van de gedragsinterventie volstaat een eenmalige onderzoeksgroep niet, maar is voortdurende meting van behandelresultaten van deelnemers aan de interventie noodzakelijk.
- Thans schrijft criterium 10 voor dat een procesevaluatie vooraf dient te zijn gegaan aan een effectevaluatie. De gedachte is namelijk dat het pas zin heeft om het effect van de interventie te meten als duidelijk is dat deze conform het protocol wordt uitgevoerd. Nadeel is echter dat het langer gaat duren voordat met het effectiviteitsonderzoek kan worden begonnen.

Conclusie

Indieners melden dat het voldoen aan de vereisten zoals geformuleerd onder criterium 10 problematisch is. Deze ervaringen worden deels weerspiegeld in de beoordelingen door de Erkenningscommissie: er wordt verhoudingsgewijs laag gescoord op dit criterium.

Tegenover deze bevinding staat het relatief geringe aantal nulcores op criterium 10 en het feit dat een erkenning nooit is geblokkeerd door een onvoldoende onderzoeksdesign.

⁹ Volgens de huidige systematiek is het mogelijk dat een potentieel effectieve gedragsinterventie niet kan worden erkend, omdat het onderzoeksdesign van onvoldoende kwaliteit is. Deze situatie heeft zich, zoals aangegeven in de vorige paragraaf, nog niet voorgedaan.

Het probleem met criterium 10 ligt voor een belangrijk deel in de bruikbaarheid en toepasbaarheid van het beschreven onderzoeksdesign. Naast de ervaringsfeiten is een aantal andere inhoudelijke kanttekeningen te plaatsen bij criterium 10.

Op basis van bovenstaande ervaringen en kanttekeningen is de Erkenningscommissie van oordeel dat het huidige criterium 10 aanpassing behoeft.

4 Voorstel wijziging criterium 10

Deze paragraaf beschrijft een voorstel voor wijziging van criterium 10. Eerst zal de voorgestelde nieuwe inhoud van criterium 10 worden weergegeven. Daarna wordt ingegaan op de onderbouwing voor de voorgestelde wijziging.

4.1

Nieuw criterium 10

Criterium 10 schrijft voor dat er moet zijn voorzien in:

1. Een Routine Outcome Monitoring (ROM).

De commissie wil bij de ex ante beoordeling van gedragsinterventies in ieder geval een plan zien waarin wordt aangegeven hoe een systeem van meten en hermeten wordt vormgegeven. Het betreft het voor, tijdens en na de behandeling verzamelen van relevante uitkomstgegevens over de toestand van de justitiabele, met behulp van valide en betrouwbare instrumenten. De gegevens bieden daarmee informatie over de doeltreffendheid van de behandeling. De eisen ten aanzien van een dergelijke ROM hebben betrekking op de te meten variabelen, de te hanteren meetinstrumenten, de meetmomenten, etc.¹⁰

Een adequaat plan voor een ROM levert 1 punt op en is een voorwaarde om een erkenning te krijgen.

2. Een onderzoeksdesign

Als de indiener daarnaast aantoonde de experimentele groep te kunnen vergelijken met een vergelijkbare controlegroep zodat het design voldoet aan de eisen die horen bij niveau 3 of hoger van de Maryland Scientific Methods Scale, dan kan op criterium 10 in totaal 2 punten worden gescoord.

4.2

Onderbouwing

Onderzoeksdesign

Ondanks de principiële en praktische bezwaren die er tegen criterium 10 gemaakt kunnen worden, is de Erkenningscommissie van mening dat dit criterium een plek moet hebben in het toetsingskader en onderdeel moet blijven van de ex ante beoordeling van de effectiviteit van de gedragsinterventie. Het uiteindelijke doel van zowel het Ministerie van Veiligheid en Justitie, de indieners en uitvoerders van gedragsinterventies, als de Erkenningscommissie is dat kan worden vastgesteld dat de gedragsinterventies die in justitieel kader worden uitgevoerd bewezen bijdragen aan het terugdringen van recidive. Het is met het oog hierop meer dan wenselijk dat het uit te voeren effectiviteitsonderzoek van meet af aan aandacht krijgt. Door het aantonen van effectiviteit op te nemen als eis voor de ex ante beoordeling van gedragsinterventies, worden

¹⁰ Na akkoord door de opdrachtgever met de wijziging van criterium 10, zal een nadere uitwerking worden gegeven aan de gestelde eisen en verdere procedure.

ontwikkelaars en indieners al vanaf het begin gestimuleerd om stil te staan bij het onderzoeksdesign en aandacht te besteden aan de wijze waarop de effectiviteit uiteindelijk kan worden aangetoond.

De Erkenningscommissie heeft echter wel oog voor de pragmatische bezwaren die enerzijds kleven aan het in een te vroegtijdig stadium opstellen en gedetailleerd uitwerken van onderzoeksplannen en anderzijds het te weinig rekening houden met wat er mogelijkwijs al bekend is over de effectiviteit van de gedragsinterventie: als er al kwalitatief goede onderzoeksresultaten bekend zijn, dan is het wellicht niet nodig nieuwe evidentie aan te tonen, of kan een minder sterk onderzoeksdesign volstaan. De methodologische en inhoudelijke eisen ten aanzien van het uit te voeren effectonderzoek spelen eigenlijk pas een doorslaggevende rol bij de ex post beoordeling.

Om bovengenoemde reden wordt geadviseerd om in het nieuwe criterium 10 minder strikte eisen te stellen aan het onderzoeksdesign. Er wordt, indien evidentie uit eerder onderzoek ontbreekt, minstens een onderzoeksdesign met controleconditie verwacht (minimaal MSMS niveau 3). Het uitvoeren van een quasi-experimenteel onderzoek met statistische controle voor specifieke variabelen (MSMS niveau 4), dan wel een experimenteel onderzoek met gerandomiseerde controlegroep (MSMS niveau 5) is niet vereist; het eventueel uitvoeren ervan kan door de indieners (of andere partijen) worden besloten. De commissie zal er zich toe beperken ex post na te gaan of er voldoende evidentie voor effectiviteit bestaat, zoals ook bij de ex ante beoordeling middels criterium 4 gebeurt. Daardoor drukt het al dan niet aanleveren van een onderzoeksdesign t.b.v. de ex ante beoordeling, minder zwaar op de totale beoordeling.

Bij kleine doelgroepen is het niet altijd mogelijk om een controlegroep samen te stellen. In dat geval kan worden uitgeweken naar een ander onderzoeksdesign.

De vereiste dat een procesevaluatie altijd vooraf moet gaan aan een effectiviteitsonderzoek, komt in het nieuwe criterium 10 te vervallen. Vanuit wetenschappelijk oogpunt is er niets op tegen om de procesevaluatie en de effectevaluatie gelijktijdig uit te voeren. Bovendien levert het gelijktijdig uitvoeren een rijkere schakering aan informatie op, omdat het (uitblijven van) effect beter gerelateerd kan worden aan de wijze waarop de interventie wordt uitgevoerd.

Routine Outcome Monitoring (ROM)

In het nieuwe criterium 10 wordt gevraagd een plan in te leveren voor een ROM. Routine Outcome Monitoring is een methodiek om, door het regelmatig verzamelen van gegevens, inzicht te krijgen in de justitiabele, de behandeling en het resultaat van de behandeling. Essentie is dat de voortgang van elke behandeling wordt gemeten, en dat de resultaten op individueel niveau en op geaggregeerd niveau beschikbaar zijn.

De Erkenningscommissie acht het van groot belang om bij de ex ante beoordeling het meten en hermeten onderdeel te laten zijn van de gedragsinterventie. Door middel van het meten en hermeten wordt de feedback op de uitvoering van de gedragsinterventie georganiseerd. Inmiddels is aangetoond dat feedback bijdraagt aan de kwaliteit van de uitvoering en daarmee aan de effectiviteit van de gedragsinterventie.

Een ROM betreft een structurele meting van de behandelingen die in het kader van de gedragsinterventie worden uitgevoerd. Dat wil zeggen dat het meten en hermeten niet alleen plaatsvindt ten behoeve van, en gedurende de uitvoering van een effectonderzoek; het meten van de voortgang van een behandeling is een essentieel onderdeel van de gedragsinterventie zelf. Een meten-hermeten systematiek (continuous monitoring) is idealiter geïntegreerd in de uitvoeringspraktijk waarbij resultaten telkens aan de behandelaar worden teruggekoppeld.

Continuous monitoring helpt om de behandeling goed uit te voeren en maakt het mogelijk om resultaten onderling te vergelijken (bijv. tussen instellingen / afdelingen / subdoelgroep). De voorgestelde verandering van criterium 10 sluit aan bij ontwikkelingen in de GGZ, waar in het kader van verantwoording (benchmarking) en onderzoek, ROM sector breed wordt opgezet. De forensische psychiatrie is hier eveneens in betrokken. Een ROM biedt tevens de mogelijkheid om de doelmatigheid van verschillende interventies met elkaar te vergelijken en kan inzicht geven in de toegevoegde waarde van een interventie in het totale aanbod van interventies. Bovendien kan een ROM, indien voldoende nauwkeurig uitgevoerd, gegevens genereren ten behoeve van een effectiviteitsonderzoek.

Puntentelling

Zoals hierboven aangegeven, wordt bij criterium 10 een ROM van grote waarde geacht. Bij de ex ante beoordeling dient in ieder geval een deugdelijk plan voor een ROM aangeleverd te worden om te worden erkend. Zonder dit plan zal 0 punt op criterium 10 worden gescoord en zal de gedragsinterventie dus niet worden erkend.

Aan een plan voor een goede ROM wordt 1 punt toegekend.

Voor criterium 10 is vervolgens 2 punten te behalen als het daarnaast een onderzoeksdesign op minimaal MSMS niveau 3 bevat.

5 Conclusie

In de inleiding is gerefereerd aan twee aandachtspunten met betrekking tot de huidige beoordelingscriteria, te weten:

- of in de huidige beoordelingscriteria de aansluiting op de uitvoeringspraktijk in Nederland in voldoende mate is meegenomen en of de criteria op dit punt aanpassing behoeven;
- of de vraag of met de huidige beoordelingscriteria voldoende rekening kan worden gehouden met gedragsinterventies voor kleine doelgroepen.

De Erkenningscommissie is verzocht om in het kader van deze twee aandachtspunten te adviseren over criterium 10.

Voor de totstandkoming van onderhavig advies is een analyse gemaakt van de scores op de beoordelingscriteria, er zijn gesprekken gevoerd met ontwikkelaars, indieners en onderzoekers, en de richting van het advies is besproken met het WODC en het Nederlands Jeugdinstituut.

Daarnaast is criterium 10 herhaaldelijk onderwerp van bespreking geweest in de Erkenningscommissie.

Met betrekking tot bovengenoemde aandachtspunten is de Erkenningscommissie van oordeel dat het huidige criterium 10 inderdaad te weinig aansluit op de uitvoeringspraktijk in Nederland. Ook houden de huidige criteria onvoldoende rekening met kleine doelgroepen.

De commissie adviseert om de inhoud van criterium 10 zodanig te wijzigen dat er minimaal een structureel systeem van meten en hermeten wordt verlangd (ROM) en daarnaast een onderzoeksdesign dat voorziet in een experimentele en controlegroep; randomisatie is niet vereist. Er kan 1 punt worden gescoord indien er een ROM wordt voorgesteld en 2 punten in geval van een ROM én een design voor (quasi-)experimenteel onderzoek.

Voor kleine doelgroepen is – onderbouwd - een afwijkend onderzoeksdesign mogelijk.

De Erkenningscommissie realiseert zich dat een deel van de genoemde knelpunten en kanttekeningen mogelijk niet geheel oplosbaar is door middel van een wijziging van criterium 10. Wel is zij van mening met het huidige advies een bijdrage te leveren aan het verbeteren van de uitvoerbaarheid van criterium 10 en daarmee aan het aanbod van bewezen effectieve interventies.

Indien het Ministerie van Veiligheid en Justitie akkoord gaat met het voorstel voor wijziging van criterium 10, dan zal de Erkenningscommissie het toetsingskader hierop aanpassen door een nadere invulling te geven van de procedures en voorwaarden. Ook kan de commissie dan terugkomen op het inrichten van effectiviteitsonderzoek bij kleine doelgroepen en op de recidivemeting bij deelnemers die lang verblijven in een gesloten residentiële setting.