

WETENSCHAP

Werkt forensische behandeling?

Een compilatie van onderzoek naar behandel-effectiviteit in de Nederlandse klinische forensische zorg

Vivienne de Vogel, Marije Keulen-de Vos, Stefan Bogaerts, Erik Bulten, Monique Delforterie, Erwin Schuringa & Paul Ter Horst

1 Inleiding

Het doel van de klinische forensische zorg, of meer specifiek de tbs-maatregel, is het beveiligen van de maatschappij; op korte termijn door iemand uit de maatschappij te halen en op langere termijn door behandeling gericht op het verlagen van risicofactoren en het opbouwen of versterken van beschermende factoren. In de media verschijnen met enige regelmaat kritische verhalen over de forensische zorg, meestal naar aanleiding van een ernstig incident, zoals een delict gepleegd door een tbs-patiënt op verlof. De vraag die daarbij steeds wordt opgeworpen, is hoe effectief de tbs-maatregel en behandeling in de forensische zorg is. Het is logisch dat er maatschappelijke onrust ontstaat bij ernstige incidenten en de opgeworpen vragen vanuit de maatschappij zijn terecht. Toch is enige nuancering hier op zijn plaats, aangezien recidive tijdens forensische behandeling uitzonderlijk is. Zo werd in een recent onderzoek met gegevens van het Adviescollege Verloftoetsing tbs (AVT) gevonden dat slechts bij 0,15% van de 15.050 positief beoordeelde verlofaanvragen sprake was van een ongeoorloofde afwezigheid met recidive.¹ Verder blijkt al jaren uit onderzoek van het Wetenschappelijk Onderzoek- en Datacentrum (WODC)² dat ernstige recidive na ontslag uit de forensische zorg, met name de tbs, relatief laag is, in ieder geval aanzienlijk lager dan na een gevangenisstraf. Hierbij dient aangetekend te worden dat de cijfers lastig te vergelijken zijn vanwege belangrijke verschillen tussen de groepen. De recidivecijfers tijdens en na forensische zorg zijn dus relatief gunstig, maar een delict kan enorme impact hebben en het zo veel mogelijk voorkomen van recidive blijft het ultieme doel van de behandeling in de forensische zorg. Het is nog onduidelijk wat precies bijdraagt aan recidivereductie en hoe behandel-effectiviteit of behandel-succes gedefinieerd kan worden.

1 Bulten et al. (2022).

2 Binnen twee jaar na tbs-behandeling: 19,2% algemene recidive: recidive ongeacht aard en ernst van gepleegd delict; 9,1% tbs-waardig: betreft een delict waarvoor opnieuw tbs kan worden opgelegd; 4,0% zeer ernstige recidive: betreft een delict met een maximale strafdreiging van acht jaar of meer. Deze cijfers zijn gebaseerd op uitstroomcohort 2011-2015. Zie voor de meest recente recidivecijfers www.wodc.nl.

Vivienne de Vogel e.a.

In 2016 publiceerden Nagtegaal en collega's³ een kosten-batenanalyse over de tbs-maatregel op basis van de beschikbare literatuur, waaruit ze concluderen dat de tbs-maatregel een waardevolle bijdrage levert aan de veiligheid van de maatschappij. In hun analyse schrijven de auteurs:

'De kosten bedragen ongeveer € 1,5 miljoen per behandeling en de recidive na behandeling is 9,3% na 2 jaar en 21,2% na 9 jaar (tbs-waardige recidive), wat lager is dan de ernstige recidive van ex-gedetineerden, die 38,9% na 2 jaar is en 63,8% na 9 jaar. Qua baten zijn de afname van psychopathologische symptomen en (overige) risicofactoren, heropnames (zowel justitieel als niet-justitieel en zowel vrijwillig als gedwongen) en de toename van beschermende factoren belangrijk. Experimenteel onderzoek is niet mogelijk. Vergelijkingen met alternatieven voor de tbs-behandeling laten zien dat de tbs-populatie aanzienlijk verschilt op belangrijke punten als (strafrechtelijk) kader, psychiatrische problematiek, delictgedrag en/of doelstelling van het verblijf, terwijl de kostprijs per dag weliswaar hoger ligt dan die voor ppc en ggz, maar niet exorbitant hoger.' (p. 743)

De conclusie van deze analyse is een hoopvolle, maar het zegt nog niet veel over de effectiviteit van de behandeling, noch geeft het inzicht in de werkzame elementen van de behandeling. Daarvoor is gedegen effectiviteitsonderzoek nodig. In de praktijk is dit echter zeer lastig uitvoerbaar, zo niet onmogelijk, zoals Nagtegaal et al. (2016) schrijven. Je kunt immers niet een controlegroep samenstellen met dezelfde typen delicten en stoornissen, die geen behandeling geven, en vervolgens bezien wat er zal gebeuren. Ook is het niet eenvoudig om goede uitkomstmaten vast te stellen en worden er nogal eens verschillende definities gehanteerd, bijvoorbeeld wat precies *de behandeling* is (zie voor nadere bespreking par. 2).

Kortom, zowel vanuit de maatschappij, als vanuit het forensische veld is er een grote behoefte aan meer kennis over effectiviteit van de geboden behandeling. Het is belangrijk om te weten of de behandeling die in de forensische zorg geboden wordt effectief is en recidive voorkomt, maar ook of die bijdraagt aan re-integratie in de maatschappij en herstel van een patiënt, bijvoorbeeld in relaties, gezondheid en kwaliteit van leven.⁴ Deze kennis kan helpen de praktijk te verbeteren en is relevant voor beleidsmakers. Het is daarnaast ook belangrijk te achterhalen wat niet werkt, zodat de behandeling zo effectief en efficiënt mogelijk ingezet kan worden en de behandelduur zo kort mogelijk gehouden kan worden. Uit een analyse van stelselwijzigingen in de forensische zorg en de bespreking van de gevolgen hiervan blijkt dat incidenten vaak sterk bepalend zijn geweest voor beleidsveranderingen (zie p. 60 *LES 8: De forensische zorg presteert steeds beter ten aanzien van recidivereductie, maar incidenten bepalen grotendeels het beeld en het beleid*⁵). Het is dan ook belangrijk om actief kennis te vergaren over de effectiviteit van forensische zorg om daarmee de behandeling verder te verbeteren en deze kennis te verspreiden,

3 Nagtegaal et al. (2016).

4 Wijma & Bulten (2021).

5 Van der Wolf et al. (2020).

ook om eenzijdige en negatieve beeldvorming in de maatschappij niet de overhand te laten krijgen (zie ook Pijler 5 van het Kwaliteitskader Forensische Zorg, 2022⁶).

1.1 Ontwikkelingen in het verbeteren van kwaliteit in de forensische zorg

Er is de afgelopen jaren veel geïnvesteerd in het vergroten van de kwaliteit van de forensische zorg. Onderzoek in het forensische veld is toegenomen en er is meer samenwerking ontstaan tussen forensische instellingen en kennisinstellingen, bijvoorbeeld door het sinds 2012 bestaande programma Kwaliteit Forensische Zorg (KFZ), waarbinnen inmiddels meer dan tweehonderd projecten zijn uitgevoerd die diverse handreikingen, protocollen en onderzoeken hebben opgeleverd. Binnen dit programma lopen ook enkele projecten naar effectiviteit van specifieke behandelprogramma's.⁷ Vanuit het KFZ-programma is in 2020 een behoefte geformuleerd aan gezamenlijk longitudinaal onderzoek en in 2021 werd een rapport opgeleverd naar thema's voor dergelijk onderzoek in de forensische zorg.⁸ Middels een klassiek Delphi-onderzoek onderzocht een groep van in totaal 64 experts welke onderzoeksthema's relevant zijn in het forensische veld. Hieruit kwamen acht thema's naar voren waarbij het thema behandel-effectiviteit als eerste prioriteit werd aange-merkt. De vragen die de auteurs bij dit thema formuleerden, waren: 1) Wat is het effect van behandelingen op het recidiverisico van (subgroepen van) forensisch psychiatrische patiënten? 2) Wat is de (kosten)effectiviteit van behandeling X in vergelijking met behandeling Y op het recidiverisico van (subgroepen van) forensisch psychiatrische patiënten? 3) Wat is de invloed van de inzet van technologische middelen in de behandeling op het recidiverisico van forensisch psychiatrische patiënten?

Vanuit het project De Lerende Praktijk⁹ van het Expertisecentrum Forensische Psychiatrie (EFP) wordt via (wetenschappelijk) onderzoek landelijk aan kwaliteitsverbetering gewerkt om zo duurzame veiligheid en verantwoorde resocialisatie van forensische patiënten naar de samenleving te kunnen garanderen. Op een strategische conferentie in 2021 met onderzoekers vanuit alle forensische klinieken in Nederland is een meerjarenonderzoeksagenda opgesteld met vier thema's: 1) effectiviteit van behandeling; 2) responsiviteit; 3) problematisch middelengebruik; en 4) mentale veerkracht en gezond werkklimaat. De huidige compilatie is een eerste resultaat van de werkgroep met het thema effectiviteit van behandeling.

1.2 Doel van de huidige compilatie

Het doel van deze compilatie is te evalueren wat er sinds 2005 is opgeleverd aan onderzoek specifiek naar de effectiviteit van de klinische forensische zorg in Nederland, wat we daaruit kunnen leren en wat voor aanknopingspunten dit biedt

6 <https://kwaliteitskaderfz.nl>.

7 Zo zijn er projecten naar eHealth-toepassingen: <https://kfz.nl/projecten/de-effectiviteitspuzzel-van-ehealth>, <https://kfz.nl/projecten/effectiviteitsstudie-grip-app> en naar *Seeking Safety*, een cognitief gedragstherapeutische interventie voor de behandeling van de combinatie verslaving en traumagerelateerde stoornissen, specifiek bij patiënten met een licht verstandelijke beperking (LVB), zie <https://kfz.nl/projecten/effectiviteitsstudie-seeking-safety-lvb>.

8 Bongers et al. (2021).

9 <https://efp.nl/projecten/de-lerende-praktijk>.

Vivienne de Vogel e.a.

voor de toekomst. Met deze resultaten beoogt de werkgroep vervolgstappen te zetten, waaronder het opzetten van een visiedocument en gezamenlijk effectiviteitsonderzoek. Het uiteindelijke doel is de behandelpraktijk verder te verbeteren en concrete handvatten te geven aan behandelaars. Eerst beschrijven we waarom onderzoek naar effectiviteit zo ingewikkeld is om uit te voeren en gaan we nader in op internationaal onderzoek naar behandel-effectiviteit.

2 De complexiteit van onderzoek naar effectiviteit van forensische behandeling

Het uitvoeren van gedegen effectiviteitsonderzoek is niet eenvoudig.¹⁰ Hieronder worden de belangrijkste redenen hiervoor samengevat.

2.1 Controlegroep

Het gebrek aan een goede controlegroep om resultaten te vergelijken met de behandelde groep is een belangrijke belemmering voor het uitvoeren van effectiviteitsonderzoek in de forensische zorg. Het stuit uiteraard op ethische bezwaren om een groep die ernstige delicten heeft gepleegd en die behandeling behoeft deze niet te geven en te bezien wat er zal gebeuren ten opzichte van de groep die wel behandeling krijgt. Er zijn alternatieven voor een controlegroep, zoals een wachtlijstpopulatie of gevangenispopulatie. De gevangenispopulatie als geheel is echter, zoals ook Nagtegaal et al.¹¹ beschrijven, niet goed te vergelijken met de forensische-zorgpopulatie, bijvoorbeeld wat betreft kenmerken van psychische problematiek, type delict, enzovoort. Dergelijk onderzoek zou onder andere vereisen dat er vooraf wordt gematched op relevante kenmerken en dat de behandelduur en de duur van de gevangenisstraf overeenkomen. Een wachtlijstpopulatie stuit onder andere ook op ethische bezwaren, omdat je patiënten nog niet behandelt die dat wel nodig hebben, wat tot schade kan leiden bij de patiënt en/of mogelijk tot agressie naar anderen.

2.2 Onderzoeksgroep

Er zijn meerdere aspecten van de te onderzoeken groep die van invloed zijn op effectiviteitsonderzoek. Ten eerste kan er sprake zijn van selectiebias als er toestemming van de patiënt moet worden verkregen voor deelname aan het onderzoek. De ervaring leert dat niet alle patiënten willen meewerken aan wetenschappelijk onderzoek. Het is de vraag of de patiënten die geen toestemming geven op relevante kenmerken afwijken van de patiënten die dit wel doen. Zijn dit mogelijk de patiënten die minder goed reageren op de behandeling? Ten tweede is de grootte van forensische onderzoeksgroepen vaak beperkt en is er regelmatig sprake van vroegtijdige uitval. Dit is met name lastig voor longitudinaal onderzoek, waarbij patiënten langere tijd worden gevolgd en er meerdere metingen plaatsvinden. Tot slot is er in klinisch forensische populaties vaak sprake van een hoge mate van comorbiditeit

¹⁰ Harte (2019).

¹¹ Nagtegaal et al. (2016).

en een grote verscheidenheid in delictgeschiedenis, waardoor het lastig is om binnen een onderzoeksgroep te kunnen differentiëren, bijvoorbeeld naar sekse, type delict of specifieke problematiek zoals psychopathie, verslaving of licht verstandelijke beperking (LVB).

2.3 Inhoud forensische behandeling

Behandeling in de klinische forensische zorg is veelomvattend en nauwelijks in losse elementen op te knippen. Forensische behandeling is immers breder dan een bepaald behandelprogramma of psychotherapie. De gehele context, het leefklimaat en de relatie met de diverse professionals zijn ook van invloed. Daarnaast kan de behandeling niet worden beschouwd als een geïsoleerde gebeurtenis. Er kan nog van alles gebeuren in het leven van een patiënt dat van invloed is op de (behandeling en behandelbaarheid van de) patiënt, bijvoorbeeld gebeurtenissen in het sociaal netwerk.

2.4 Uitkomstmaat

Het is belangrijk om een goede, betrouwbare uitkomstmaat te hebben voor effectiviteit van de behandeling. De discussies hierover lopen uiteen en onderzoekers houden vaak verschillende definities aan. Recidive is uiteraard een belangrijke uitkomstmaat, aangezien het primaire doel van de forensische zorg is terugval te voorkomen, maar er zijn ook nadelen aan het gebruiken van recidive als uitkomstmaat. Niet alle recidive wordt ontdekt en gerapporteerd (het zogenoemde *dark number*), en er is een (relatief lange) follow-up-tijd voor nodig om recidive (als veroordeling) te kunnen detecteren, omdat er tijd zit tussen het plegen van het delict en de veroordeling ervoor. Bovendien worden verschillende definities gehanteerd, bijvoorbeeld welk type recidive het meest relevant is (speciaal, specifiek, algemeen) en naar welke cijfers je kijkt (politiegegevens, formele veroordelingen). Verder zijn andere uitkomstmaten relevant, zoals behandelvoortgang, ervaren kwaliteit van leven en herstel van een patiënt, maar deze zijn eveneens ingewikkeld te identificeren en te definiëren.¹²

Kortom, er bestaan meerdere uitdagingen in het uitvoeren van effectiviteitsonderzoek in het klinisch forensische veld. Doorgaans worden *Randomized Controlled Trials* (RCT's) als gouden standaard gezien in effectiviteitsonderzoek. Een RCT is een gecontroleerd onderzoek waarbij proefpersonen aselekt over minstens twee groepen (de behandelgroep en de controlegroep) worden verdeeld om externe variabelen evenredig te verdelen en systematische verschillen tegen te gaan (randomisatie). Daarbij worden alle variabelen, behalve de onafhankelijke variabele (de groep), zo veel mogelijk gecontroleerd of constant gehouden, zodat ze de afhankelijke variabele (de behandeling) niet beïnvloeden. In de praktijk kent dit type onderzoek vele complicaties. Hoewel RCT's als gouden standaard gelden, is er ook

12 In een lopend KFZ-project naar de operationalisatie van behandelvoortgang in de forensische zorg worden vanuit de theorieën *Risk Need Responsivity* (recidive) en *Good Lives Model* (herstel) meerdere factoren van belang onderscheiden, zie <https://kfz.nl/projecten/literatuuronderzoek-en-operationalisatie-van-behandelvoortgang-binnen-de-forensische-zorg>.

Vivienne de Vogel e.a.

kritiek op.¹³ Zo bespreekt Harte in haar oratie dat door een dergelijke opzet vooral gefocust wordt op een specifieke kortdurende interventie uitgevoerd in een geïsoleerde omgeving in plaats van de gehele behandeling en met aandacht voor de behandelrelatie.

3 Internationaal onderzoek naar behandel-effectiviteit in de forensische zorg

Recentelijk zijn er drie internationale meta-analyses verschenen naar de effectiviteit van behandeling in het forensische werkveld. Papalia et al. (2019) voerden een meta-analyse uit naar behandel-effectiviteit voor justitiabelen met geweldsdelicten in zowel justitiële als forensische klinische en ambulante instellingen in het voorkomen van recidive en wangedrag in die instellingen. Zij includeerden 27 studies met in totaal 7062 gewelddadige delinquenten (waarvan 22% vanuit forensische klinieken), en vonden een positief resultaat voor behandeling. De resultaten lieten zien dat, voor deze totale groep, de kans op zowel gewelddadige als algemene recidive kleiner was voor behandelde gewelddelinqenten dan voor onbehandelde gewelddelinqenten. Het gemiddelde effect voor gewelddadig en algemeen wangedrag in de instellingen was echter niet significant. De auteurs keken ook naar welke variabelen van invloed waren en vonden dat hoe langduriger en intenser de behandeling was, des te beter de resultaten. Ook vonden zij aanwijzingen dat behandeling binnen een therapeutische omgeving waar direct met het nieuw geleerde gedrag geoefend kan worden tot betere resultaten leidt. Ze noemen de bevindingen veelbelovend en geven aan dat multimodale¹⁴ behandelingen het meest effectief lijken. Intensieve multimodale, cognitieve gedragstherapeutische programma's die terugvalpreventie, rollenspellen en huiswerkopdrachten bevatten, en interventies gericht op versterken van interpersoonlijke vaardigheden en woedecontrole bleken gerelateerd te zijn aan sterkere reductie in gewelddadige recidive. Papalia en collega's¹⁵ maken de kanttekening dat het bestaande bewijsmateriaal beperkt is, gezien de matige kwaliteit van de meeste studies. Zij concluderen tevens dat er meer onderzoek nodig is om de behandel-effectiviteit te onderzoeken om te bepalen welke componenten effectief zijn, in welke combinatie, en voor welke personen precies. McIntosh en collega's¹⁶ voerden een meta-analyse uit naar studies in forensische instellingen, waarbij ze keken naar diverse uitkomstmaten, zoals verbetering in zelfinzicht, psychiatrische symptomen en impulsiviteit. De methodologische kwaliteit van de studies werd vooraf beoordeeld. Ze includeerden in totaal 28 studies,¹⁷ waarvan drie werden beoordeeld als van hoge kwaliteit, elf als acceptabel en veertien van lage kwaliteit. De auteurs vonden kleine effectgroottes voor behandeling ten opzichte van de vergelijkingsconditie (*Treatment As Usual*, TAU) wat betreft

13 Gerstman (2023); Harte (2019).

14 Multimodale behandeling verwijst naar de inzet van een combinatie aan behandelingen/behandel-vormen.

15 Papalia et al. (2019).

16 McIntosh et al. (2021).

17 Twee geïncludeerde studies waren ook bij de meta-analyse van Papalia et al. (2019) geïncludeerd.

verbetering van inzicht in psychische stoornissen, verbetering van psychiatrische symptomen, verbetering van probleemoplossend vermogen, vermindering van procriminele attitudes en verbetering van gedrag op de afdelingen. Verder werd een middelgroot effect gevonden voor behandeling gericht op psycho-educatie. De resultaten toonden verder aan dat behandeling geen voordeel had ten opzichte van de vergelijkingsconditie op diverse relevante kenmerken, zoals impulsiviteit, empathie, copingvaardigheden, woede of intramuraal geweld. De auteurs concluderen net als Papalia et al. dat de methodologische kwaliteit van de studies over het algemeen problematisch was en dat er meer goed opgezette onderzoeksprojecten nodig zijn om de effectiviteit van behandeling te onderzoeken.

Abbing et al.¹⁸ voerden een systematische review, narratieve synthese en meta-analyse uit naar effectiviteit van *arts therapies* (bijv. beeldende therapie, dramatherapie, dans- en bewegingstherapie) in de forensische zorg en gevangenissen. Ze concludeerden op basis van een analyse van 23 studies¹⁹ dat er significante verbeteringen werden gevonden op zowel risicofactoren (psychiatrische symptomen en verslaving) als beschermende factoren (sociaal en psychologisch functioneren). Er werd in de meeste studies helaas niet gekeken naar verbetering op antisociaal of delinquent gedrag en wanneer dit wel werd meegenomen, waren de resultaten niet significant.

Kortom, internationaal onderzoek tot nu toe biedt enige aanwijzingen voor verbeteringen op risicofactoren en vermindering van zowel gewelddadige als algemene recidive na behandeling bij volwassen (gewelddadige) daders. Hoe, waarom en voor wie de behandeling precies effectief is, blijft echter onzeker. Bovendien valt op dat de kwaliteit van de in de meta-analyses geïncludeerde studies nog behoorlijk te wensen over laat.

4 Methode

4.1 Procedure

In 2021 is vanuit het EFP een uitvraag gedaan door (onderzoekers van) alle forensische klinieken aan te schrijven en te vragen naar hun studies over effectiviteit van behandeling verschenen vanaf 2005. Er werden 69 stukken aangeleverd vanuit in totaal 11 forensische instellingen die waren verschenen in de periode 2005-2023.²⁰ Daarnaast is door de auteurs een aanvullende literatuuresearch gedaan, onder andere aan de hand van de referentielijsten van de aangeleverde stukken, waarbij nog een studie werd gevonden. Van deze zeventig stukken werd eerst een selectie gemaakt of het daadwerkelijk een studie betrof of dat het een meer algemeen of theoretisch artikel of hoofdstuk betrof. In totaal waren er 55 studies waarvan een overzicht werd gemaakt naar voorbeeld van de lijst van Papalia et al. met kenmerken van de studies, onder andere het doel en type onderzoek, de gebruikte uitkomstmaat, en of er over de resultaten was gepubliceerd. Op basis hiervan werd

18 Abbing et al. (2023).

19 Er was geen overlap met studies geïncludeerd in Papalia et al. (2019) en McIntosh et al. (2021).

20 Sommige artikelen waren in 2021 online beschikbaar, maar werden pas na 2021 officieel gepubliceerd.

Vivienne de Vogel e.a.

een inhoudelijke selectie gemaakt door de eerste twee auteurs. Studies werden niet in de verdere analyse opgenomen wanneer ze op geen enkele manier effectiviteit van behandeling onderzochten, enkel bedoeld waren als een intern rapport en niet waren gepubliceerd, niet over de klinische forensische populatie gingen, of niet over patiënten, maar over professionals gingen.

Verder was er sprake van overlap in onderzochte populaties in de publicaties. Zo waren er bijvoorbeeld meerdere publicaties over het multicenter onderzoeksproject van Bernstein et al.²¹ en meerdere studies over een multicenter HKT-R databestand.²² Er is voor gekozen hier enkel het hoofdartikel dan wel het meest recente artikel of de grootste onderzoeksgroep te beschrijven. Vervolgens werden de overgebleven studies ingedeeld in type onderzoek, te weten RCT's of andere typen onderzoek. Van alle geselecteerde gepubliceerde RCT's werd de methodologische kwaliteit beoordeeld aan de hand van de *Mixed Method Appraisal Tool Version 2018* (MMAT²³) door de eerste twee auteurs.

4.2 Instrument

De MMAT bestaat uit kwaliteitscriteria voor verschillende typen onderzoek. Het instrument is ontwikkeld in Canada en wordt regelmatig gebruikt in meta-analyses en reviews. In de huidige beoordeling is gebruik gemaakt van onderdeel 2 van de MMAT, genaamd de *Quantitative randomized controlled trials* (zie Tabel 1). De eerste twee auteurs vulden onafhankelijk van elkaar de MMAT in voor de gevonden RCT's, waarbij nagenoeg perfecte overeenstemming was.

5 Resultaten

Van de 55 in het overzicht²⁴ opgenomen studies betrof een deel ongepubliceerde stukken, zoals interne rapportages (n = 3), eindrapport voor de opleiding tot klinisch psycholoog (n = 1), of masterscripties (n = 10) die weliswaar zijn gepubliceerd via universiteiten, maar niet in een tijdschrift. Van alle in een tijdschrift gepubliceerde artikelen waren er vier in het Nederlands en 37 in het Engels. Al deze publicaties konden worden beschouwd als *peer-reviewed*, oftewel op inhoud getoetst door experts.

Bij nadere bestudering van de studies bleek een aanzienlijk deel van de aangeleverde studies niet specifiek als doel te hebben om de effectiviteit van behandeling te meten, maar er werd bijvoorbeeld onderzocht wat voorspellers zijn van incidenten tijdens de behandeling of wat het effect is van een bepaalde training op behandelaren. Een deel van deze studies was desondanks interessant om mee te nemen in de huidige compilatie. Deze worden in een aparte paragraaf beschreven. Uiteindelijk waren er zes gepubliceerde, *peer-reviewed* RCT's naar effectiviteit van behandeling die werden beoordeeld aan de hand van de methodologische kwaliteitscriteria van de MMAT (zie Tabel 1). Hieruit bleek dat de zes studies op de meeste criteria

21 Bernstein et al. (2023).

22 Janković et al. (2021).

23 Hong et al. (2018).

24 Het overzicht kan desgewenst worden opgevraagd bij de eerste auteur.

voldeden, met uitzondering van het criterium of de onderzoekers blind waren voor de conditie van de patiënt (behandeling of controleconditie). Meestal wisten de onderzoekers wel welke patiënten aan de behandeling meededen en welke TAU volgden. In de praktijk waarbij men – ook voor de veiligheid – bekend dient te zijn met de patiënten en waarbij onderzoekers nauw samenwerken met behandelaars om data te kunnen verkrijgen, is het nauwelijks mogelijk om de onderzoekers blind te houden voor de behandelconditie.

Tabel 1 MMAT criteria voor de zes RCT's

Criterion	Beoordeling
1 Is de randomisatie correct uitgevoerd?	In de meeste RCT's werd dit correct gedaan, er werden verschillende methoden of technieken gebruikt.
2 Zijn de groepen in de basis vergelijkbaar?	Ja, behalve Klein Tuente et al. (2020), waar men enkele verschillen vond in o.a. leeftijd en opleidingsniveau, waarvoor werd gecorrigeerd in de analyses.
3 Zijn de uitkomstdata volledig?	In alle RCT's was sprake van onvolledige uitkomstdata, variërend van 11,3% tot 28,9%. MMAT noemt dat in de literatuur uit wordt gegaan van drop-out tussen 5 en 20%. Twee RCT's hadden een hogere drop-out: Bernstein et al. (2023) 28,9% en Zwets et al. (2016) 27%.
4 Waren de onderzoekers die de uitkomsten maten blind voor de conditie van de groep?	Bij geen van de zes RCT's was dit het geval. Gezien de aard van de instelling en de kleinschaligheid van de studies bleek dit in de praktijk niet mogelijk.
5 Hielden de deelnemers de interventie vol?	Bij drie RCT's studies werd gerapporteerd dat dit het geval was. Bij drie RCT's werd wel de indruk gewekt dat deelnemers het hadden afgerond, maar werd dit niet expliciet vermeld in het artikel.

5.1 *Randomized Controlled Trials (RCT's)*

In Tabel 2 worden de RCT's samengevat. Er kan een onderscheid worden gemaakt naar RCT's uitgevoerd in meerdere instellingen tegelijk²⁵ en kleinschalige of pilot RCT's²⁶ die binnen één instelling plaats hadden gevonden.

5.1.1 *Multicenter RCT's*

Het meest omvangrijke en langstlopende multicenter RCT-onderzoek is van Bernstein en collega's²⁷ naar schematherapie aangepast voor forensische patiënten. Schematherapie richt zich op onvervulde emotionele basisbehoeften in de kindertijd (bijv. de behoefte aan veiligheid en stabiliteit), wat heeft geleid tot de ontwikkeling van disfunctionele schema's of gedachtenpatronen, en disfunctionele sche-

25 Bernstein et al. (2023); Klein Tuente et al. (2020).

26 Van Alphen et al. (2019); Bloemers & De Vogel (2020); Fielenbach et al. (2018); Zwets et al. (2016).

27 Bernstein et al. (2023).

Vivienne de Vogel e.a.

mamodi of emotionele toestanden die het gedrag dicteren. Schematherapie in de forensische zorg is erop gericht om deze patronen te herkennen en de intensiteit van disfunctionele schemamodi te verminderen, terwijl gezonde modi versterkt worden. In dit onderzoek werden 103 mannelijke patiënten met een verleden van geweld en voornamelijk cluster B-persoonlijkheidsstoornissen²⁸ uit in totaal zeven klinieken geïncludeerd. Een groep van 54 patiënten kreeg schematherapie en werd vergeleken met 49 patiënten die TAU kregen. Gedurende drie jaar werd iedere zes maanden door onderzoeksassistenten metingen uitgevoerd met diverse instrumenten, bijvoorbeeld een instrument om symptomen van persoonlijkheidsstoornissen te meten. Verder werd er gebruik gemaakt van risicotaxatie-gegevens en registraties van incidenten tijdens de behandeling en verlofmogelijkheden. Uit het onderzoek bleek dat zowel patiënten die schematherapie kregen als patiënten die TAU kregen verbeteringen lieten zien op diverse uitkomstmaten. De patiënten met schematherapie verbeterden echter sneller op symptomen van persoonlijkheidsstoornissen en kregen eerder toestemming voor verlof. Daarbij moet wel rekening worden gehouden met het feit dat schematherapie intensiever was dan TAU, namelijk in principe²⁹ twee keer per week, terwijl dit bij TAU eenmaal per week was. Verder waren er in dit onderzoek behoorlijk wat exclusiecriteria, waardoor de resultaten niet direct te generaliseren zijn naar de gehele forensisch klinische populatie.

Klein Tuinte et al.³⁰ voerden een RCT uit naar het effect van *Virtual Reality Agressie Preventie Training* (VRAPT) bij vier klinieken. VRAPT is een training waarbij tbs-patiënten verschillende realistische scenario's oefenen in een driedimensionale virtuele omgeving, zodat ze leren adequater om te gaan met hun reactieve agressie en impulsen. De training bestaat uit zestien tweewekelijkse individuele behandelsessies van zestig minuten onder leiding van een ervaren psycholoog of vaktherapeut. Er werden 128 mannelijke patiënten geïncludeerd en deze werden willekeurig toegewezen aan de VRAPT (n = 64) of de wachtlijstconditie (n = 64). Op drie momenten tijdens het onderzoek vonden er metingen plaats: bij aanvang, na afronding van de behandeling en drie maanden na afloop van de behandeling. De onderzoekers vonden geen significante verschillen op de primaire uitkomstmaten van geobserveerde en zelfgerapporteerde agressie. Op secundaire uitkomstmaten, gemeten met zelfrapportagelijsten, verbeterde de VRAPT-groep ten opzichte van de controlegroep bij de tweede meting op vijandigheid, controle over boosheid en een aspect van impulsiviteit. Deze verbeteringen waren echter niet meer aanwezig bij de follow-upmeting drie maanden na afloop van de behandeling. In dit onderzoek werd benadrukt dat het belangrijk is om de geoefende vaardigheden ook breder te trainen, dus ook buiten de therapiekamer. Dit bleek in de periode van dit onderzoek lastig te verwezenlijken in de instellingen waar het onderzoek plaatsvond.

28 Hieronder vallen de antisociale, borderline, narcistische en histrionische persoonlijkheidsstoornis. Zie voor meer informatie: www.ggzstandaarden.nl/zorgstandaarden/persoonlijkheidsstoornissen-zorgstandaard-2017/over-persoonlijkheidsstoornissen.

29 In de praktijk bleek tweemaal per week niet altijd haalbaar.

30 Klein Tuinte et al. (2020).

Tabel 2 RCT's effectiviteit behandeling^a

Auteurs	Design	Behandeling	Instelling	Grootte populatie	Populatie	Uitkomstmaten	Conclusies
Van Alphen et al., 2019	RCT-pilot	MACT	Inforsa	16 exp. 15 controle	Psychotische problematiek, 69% man	Aandacht (o.a. subtest van WAIS-IV-NL)	MACT heeft positief effect op 3 maten van aandacht.
Bernstein et al., 2023	RCT	Schema-therapie	7 klinieken	54 exp. 49 controle	Mannen, persoonlijkheidsstoornis, paranoia, cluster B	Zelfrapportagelijsten, risicotaxatie, interviews, geregistreerde incidenten, toegekend verlof	Verbetering beide groepen, behandelgroep sterker en sneller, o.a. sneller verlof, verbetering symptomen persoonlijkheidsstoornis.
Bloemers & De Vogel, 2020	RCT-pilot	COMET	de Forensische Zorgspecialisten	5 exp. 7 controle	Psychotische problematiek, paranoia 79% man	Rosenberg Self-Esteem Scale, Green Paranoid Thoughts Scale	Klein effect voor COMET in verbetering zelfbeeld, geen verbetering paranoia.
Fielenbach et al., 2018	RCT	Neurofeedback-training	Van Mesdag	21 exp. 21 controle	Mannen verslavingsproblematiek	Divers, EEG, zelfrapportagelijsten	Verbetering beide groepen voor impulsiviteit en zucht naar drugs, geen meerwaarde van neurofeedbacktraining.

Tabel 2
(Vervolg)

Auteurs	Design	Behandeling	Instelling	Grootte populatie	Populatie	Uitkomstmaten	Conclusies
Klein Tuente et al., 2020	RCT	VRAPT	4 klinieken	64 exp. 64 controle	Mannen	Observatie en zelfrapportagelijsten, interview	Geen effect voor agressie; korte termijn (3 maanden) effect op maten van vijandigheid en woede.
Zwets et al., 2016	RCT-pilot	PMT	Kijvelanden	22 exp. 15 controle	Mannen aangemeld voor ART	Observatie en zelfrapportagelijsten	Beide groepen verbeterden. Geen extra effect PMT op agressie, wel effect op doelen van PMT, namelijk verbetering lichaamsbewustzijn tijdens spanningsvolle situaties.

a Exp = experimenteel; MACT = Musical Attention Control Therapy; COMET = Competitive Memory Training; VRAPT = Virtual Reality Aggressie Preventie Training; PMT = Psychomotorie training; ART = Aggression Replacement Training. Voor de volledige informatie over bijvoorbeeld gebruikte instrumenten en uitkomstmaten verwijzen we naar de betreffende artikelen.

5.1.2 RCT's uitgevoerd binnen één instelling

Fielenbach et al.³¹ onderzochten het effect van neurofeedbacktraining bij 21 mannelijke patiënten met middelenmisbruik en hoge mate van impulsiviteit in de Van Mesdag. Neurofeedback is een methode gericht op het veranderen van hersenactiviteit, gebaseerd op het lerende vermogen van het brein. Deze patiënten kregen twintig sessies neurofeedbacktraining als aanvulling op TAU en hun resultaten op maten van impulsiviteit, zoals gemeten met zelfrapportage, en zucht naar en gebruik van drugs werden vergeleken met een controlegroep van 21 patiënten die enkel TAU kregen. Er werd in beide groepen een verlaging gevonden van zucht naar drugs en impulsiviteit, er waren geen significante verschillen tussen de twee groepen. Een kanttekening bij dit onderzoek was dat de geboden filmpjes tijdens de neurofeedbacktraining vrij simplistisch van vorm waren en door meerdere deelnemers als saai werden ervaren. De auteurs bevelen aan om het onderzoek uit te breiden met nieuwere vormen van neurofeedbacktraining. Verder viel in dit onderzoek op dat geen enkele patiënt de sessies binnen de geplande tijdsperiode van tien weken had weten te voltooien.

Van Alphen et al.³² onderzochten het effect van *Musical Attention Control Therapy* (MACT) bij zestien patiënten van Inforsa met psychotische problematiek en problemen met aandacht. De therapie bestond uit een dertig minuten durende groepstraining (twee tot zes personen) die éénmaal per week gedurende zes weken werd geboden in aanvulling op TAU. MACT is een geprotocolleerde techniek gericht op hersenactiviteit tijdens muziek maken of beluisteren om aandachtsprocessen te optimaliseren. In de sessies werd naast de muzikale oefeningen ook psycho-educatie geboden. De onderzoekers vonden dat de experimentele groep significant meer verbeteringen liet zien op verschillende maten van aandacht in vergelijking met de controlegroep van vijftien patiënten die TAU ontvingen en die op een wachtlijst stonden voor MACT. De auteurs maken hierbij de kanttekening dat de positieve effecten mogelijk ook deels te verklaren zijn door het feit dat deze patiënten extra aandacht kregen. De auteurs bevelen aan nader onderzoek te doen in grotere groepen en met meer specifieke MACT-interventies.

In een kleine pilotstudie bij de Forensische Zorgspecialisten werd bij forensische patiënten met paranoïde problematiek onderzocht of de *Competitive Memory Training* (COMET), een cognitieve gedragstherapeutische interventie bedoeld om het zelfbeeld te verbeteren, leidde tot verbeteringen in zelfbeeld en reductie van paranoia.³³ De COMET-interventie werd vergeleken met TAU. Metingen met zelfrapportagelijsten voor zelfbeeld en paranoia werden uitgevoerd bij de start, na tien weken (*post treatment*) en acht weken follow-up, bij in totaal vijf patiënten in de experimentele conditie en zeven in de controleconditie. Post-treatmentresultaten lieten een significante verbetering zien op zelfbeeld bij de vijf patiënten die COMET hadden gevolgd, maar geen reductie van paranoia lieten zien. Het viel verder op dat, hoewel patiënten positief waren over de geboden behandeling en ze trouw op de afspraken kwamen en de behandeling hadden afgerond, ze in de praktijk

31 Fielenbach et al. (2018).

32 Van Alphen et al. (2019).

33 Bloemers & De Vogel (2020).

Vivienne de Vogel e.a.

moeite hadden met het voltooien van huiswerkopdrachten. De auteurs bevelen aan om COMET verder aan te passen voor de forensische doelgroep, waarbij rekening wordt gehouden met de beperkte belastbaarheid en dat er meer ondersteuning en structuur wordt geboden bij het uitvoeren van huiswerkopdrachten.

Zwets et al.³⁴ voerden een RCT uit naar het effect van psychomotore therapie als aanvulling op *Aggression Replacement Training* (ART) bij de Kijvelanden. De psychomotore therapie (PMT) was gericht op het beter herkennen van situaties van boosheid, het beter herkennen van lichamelijke sensaties bij boosheid (lichaamsbewustzijn), acceptatie van boosheid als emotionele reactie, en het verbeteren van vaardigheden om boosheid en agressie te reguleren. PMT bestond uit 25 wekelijkse sessies van 90 minuten. Er werden 37 mannelijke patiënten geïncludeerd en deze werden willekeurig toegewezen aan de groep PMT + ART (n = 22) of aan de controlegroep (n = 16) die PMT + reguliere sportmomenten kreeg aangeboden. Op drie momenten tijdens het onderzoek vonden er metingen plaats: bij aanvang, na afronding van de behandeling en vijftien weken na afloop van de behandeling. De onderzoekers vonden geen significante verschillen op de primaire uitkomstmaten van geobserveerde en zelfgerapporteerde agressie. Op secundaire uitkomstmaten, gemeten met zelfrapportagevragenlijsten, verbeterde de PMT + ART-groep ten opzichte van de controlegroep wat betreft lichaamsbewustzijn tijdens spanningsvolle situaties. De auteurs noemen als kanttekeningen bij deze RCT het hoge percentage drop-out (27%) en de inzet van zelfrapportagevragenlijsten (met name t.a.v. de secundaire uitkomstmaten), en de lage *base rate*³⁵ van agressie als gevolg van de restrictieve setting.

5.2 Onderzoek naar verandering in risico en beschermende factoren

Meerdere van de door de klinieken aangeleverde studies hadden tot doel het meten van verandering in dynamische risico- en of beschermende factoren met de standaard gebruikte risicotaxatie-instrumenten. Ondanks dat deze studies niet specifiek effectiviteit van behandeling onderzochten en dit ook niet als variabele hadden meegenomen, geven de resultaten wel een indicatie van de behandeling als geheel. Immers, forensische behandeling is gericht op het verminderen van risicofactoren en het versterken van beschermende factoren.

In deze studies werden positieve veranderingen gevonden, dat wil zeggen een significante verlaging van het recidiverisico zoals gemeten met de risicotaxatie-instrumenten voor de groep als geheel. Jankovic et al.³⁶ onderzochten veranderingen op dynamische risicofactoren, zoals gemeten met de HKT-R op vijf momenten van 722 mannelijke patiënten die tussen 2004 en 2014 onvoorwaardelijk waren ontslagen uit twaalf klinische forensische instellingen. Er werd een significante verlaging van dynamische risicofactoren gevonden. Tussen opname en het eerste ongebegeide verlof vonden de grootste veranderingen plaats en dit eerste ongebegeide verlof werd door de auteurs dan ook beschreven als een *turning point*. Er waren wel ver-

34 Zwets et al. (2016).

35 *Base rates* zijn statistieken die worden gebruikt om het percentage van een populatie te beschrijven dat een bepaald kenmerk vertoont.

36 Jankovic et al. (2021).

schillen; patiënten met gediagnosticeerde verslavingsproblematiek en psychotische problematiek veranderden het meest in de eerste fase van de behandeling, terwijl patiënten met een cluster B-persoonlijkheidsstoornis juist de meeste verandering lieten zien in een latere fase van de behandeling. In de studie van De Vries Robbé et al.,³⁷ waarin een groep van 108 mannelijke patiënten uit de Van der Hoeven Kliniek en de Van Mesdag Kliniek werd meegenomen, werd naast verlaging van risicofactoren gemeten met de HCR-20 Versie 2 tevens verhoging van beschermende factoren gevonden, zoals gemeten met de SAPROF, en bleek dat positieve veranderingen significant gerelateerd waren aan minder recidive.

Hoewel uit de bovenbeschreven studies geen directe conclusies te trekken zijn over effectiviteit van behandeling, wijzen de bevindingen van verlaging van dynamische risicofactoren op zowel de HKT-R als de HCR-20 en verbeteringen op beschermende factoren met de SAPROF, op algemeen slagen van de behandeling in het verlagen van het recidiverisico. Toekomstig onderzoek zal moeten kijken naar meerdere uitkomstmaten op langere termijn, zoals recidive na ontslag en herstel op verschillende levensdomeinen.

5.3 Overige onderzoeken

In een onderzoek van Smid et al.³⁸ werden recidivecijfers vergeleken van 90 zedendelinquenten die een tbs-behandeling hadden ondergaan en 176 zedendelinquenten die alleen in de gevangenis hadden gezeten. De groepen werden gematcht op risiconiveau, zoals gemeten met het risicotaxatie-instrument de STATIC-99R. De gemiddelde follow-upperiode was 148 maanden. De resultaten lieten zien dat tbs'ers minder recidiveerden dan onbehandelde plegers, vooral wanneer gekeken werd naar gewelds- en/of zedenrecidive. Zo viel 52% van de tbs'ers met een hoog recidiverisico terug, terwijl van dezelfde risicogroep de onbehandelde delinquenten in detentie 77% terugviel. Het verschil tussen de groepen gold alleen voor de plegers met een matig-hoog tot hoog recidiverisico. Plegers met een laag tot laag-matig risico recidiveerden aanzienlijk minder en een tbs-behandeling maakte daarin geen opvallend verschil. De auteurs merken op dat het in dit onderzoek ging om een groep die al lang geleden is uitgestroomd en dat er reden is om aan te nemen dat de tbs-behandeling er sinds die tijd nog beter op is geworden.

Verder zijn er diverse onderzoeken gedaan met als doel het voorspellen van recidive of van incidenten tijdens de behandeling of instrumentontwikkeling om verandering te meten,³⁹ of studies naar hoe patiënten zelf de behandeling ervaren. Deze studies zeggen niet direct iets over het effect van de behandeling, maar laten wel zien welke onderdelen belangrijk zijn en waar behandeling op gericht zou moeten zijn. Vermeldenswaardig is tot slot dat er meerdere effectiviteitsstudies zijn uitgevoerd naar behandeling bij patiënten met een licht verstandelijke beperking (LVB), opgenomen in Trajectum. Deze studies includeerden echter een bredere populatie dan enkel patiënten met een forensische zorgtitel. De uitkomsten voor de LVB-populatie waren over het algemeen veelbelovend, bijvoorbeeld in een longitudinale

37 De Vries Robbé et al. (2014).

38 Smid et al. (2016).

39 Schuringa et al. (2019).

Vivienne de Vogel e.a.

studie naar *Flexible Assertive Community Treatment* (FACT), waarbij significante verbeteringen op sociaal en psychologisch functioneren werden gevonden.⁴⁰

6 Discussie

In de huidige compilatie werd het onderzoek naar behandel-effectiviteit in de klinische forensische zorg gepubliceerd in de periode 2005-2023 besproken. Ondanks dat er in totaal zeventig stukken waren aangeleverd dan wel gevonden, waren er maar zes RCT's die daadwerkelijk als doel hadden de behandel-effectiviteit te meten. Dat er relatief weinig RCT's zijn uitgevoerd, is niet verwonderlijk, gezien de complexiteit van het klinische forensische veld en de in de inleiding beschreven uitdagingen, zoals het gebrek aan een goede controlegroep. Deze zes RCT's voldeden voldoende aan de algemeen aanvaarde kwaliteitseisen voor effectiviteitsonderzoek, met uitzondering van het criterium dat onderzoekers blind zouden moeten zijn voor de conditie.

6.1 Conclusies uit de RCT's

Uit de zes RCT's kan, hoewel er enkele kleine positieve of kortetermijn-behandel-effecten worden gerapporteerd, niet worden geconcludeerd dat de onderzochte interventies onmiskenbaar een groter effect op de uitkomstmaat hadden dan de TAU. De voornaamste redenen hiervoor zijn de te kleine steekproeven en de statistische keuzes (zie verder hieronder). Voor schematherapie, MACT en COMET werden significante effecten gerapporteerd; de experimentele groepen verbeterden sterker en/of sneller dan de controlegroepen. Voor VRAPT werd enkel een kortetermijneffect gevonden op maten van vijandigheid en woede en voor PMT een klein effect voor verbetering van lichaamsbewustzijn in spanningsvolle situaties. Voor neurofeedbacktraining werd geen extra effect gevonden ten opzichte van TAU. Positief is dat er ook bij de patiënten die TAU ontvingen of die in de vergelijkingsconditie zaten positieve veranderingen werden gevonden.⁴¹

Alle RCT's hadden belangrijke beperkingen. Zo waren de onderzoekers die de uitkomstmaten afnamen nooit blind voor de conditie, wat in de praktijk ook zeer lastig uitvoerbaar is. De groepen waren meestal te klein en selectief, specifiek gericht op bepaalde stoornissen en meestal werden alleen mannelijke patiënten geïncledeerd. De resultaten zijn dan ook moeilijk te generaliseren naar de forensische klinische populatie als geheel. Over de behandeling van de relatief kleine, maar groeiende groep vrouwen in de forensische zorg weten we weinig; daar is meer onderzoek voor nodig.⁴² Verder waren er tekortkomingen in de gebruikte statistiek; zo werd er doorgaans bijvoorbeeld niet gecorrigeerd voor multiple testen, en werden doorgaans geen powerberekeningen gedaan. Ook waren er beperkingen wat betreft de uitkomstmaten. Soms werd enkel gebruik gemaakt van zelfrapportagelijsten, en in geen van de RCT's werden formele recidivedata gebruikt, omdat dat niet mogelijk was door de relatief korte follow-up tijden. Voor toekomstig onderzoek is het van

40 Neijmeijer et al. (2019).

41 Bernstein et al. (2023); Fielenbach et al. (2018); Zwets et al. (2016).

42 Zie ook Papalia et al. (2019); De Vogel et al. (2023).

belang langere follow-up-tijden te hanteren en gebruik te maken van diverse uitkomstmaten, om ook te onderzoeken of de effecten na langere tijd beklijven. Zo werd in de RCT van Klein Tuente et al.⁴³ bijvoorbeeld wel direct na behandeling een significant effect gevonden, maar was dit na drie maanden verdwenen. Wat is nodig om met het geleerde te blijven oefenen zodat het beklijft?

Een belangrijke beperking van RCT-onderzoek werd besproken door Van Alphen et al.,⁴⁴ namelijk de extra tijd en aandacht die de experimentele groep krijgt omdat de onderzochte behandeling vaak als aanvulling op TAU werd geboden. Dit geldt tevens als mogelijke beperking bij de andere RCT's. Meer in het algemeen is gevonden dat des te langduriger en intenser de behandeling is, des te beter de resultaten zijn.⁴⁵ Bovendien kan hierbij een rol spelen dat therapeuten extra training kregen en wellicht extra gemotiveerd of enthousiast waren over de experimentele behandelvorm, wat mogelijk een gunstig effect heeft op de resultaten. Dit geldt overigens ook voor de betrokken onderzoekers.⁴⁶ Tot slot waren er beperkingen aan de geboden behandelingen, zoals besproken in de verschillende RCT's. Sommige behandelingen staan nog redelijk in de kinderschoenen of waren (nog) niet optimaal vormgegeven voor de forensische populatie. Meer in het algemeen speelt mee dat er tegelijk met de experimentele behandelvorm nog allerlei andere variabelen van invloed kunnen zijn waarvoor niet altijd gecontroleerd kan worden, bijvoorbeeld de sterk gestructureerde en gecontroleerde omgeving in de instelling en medicatiegebruik.

Daarnaast waren er meerdere studies die, ondanks dat ze niet direct gericht waren op het meten van effectiviteit van behandeling, interessante inzichten en handvatten voor de dagelijkse klinische forensische praktijk opleverden. Aanwijzingen voor positieve effecten van de forensische behandelingen werden gevonden in meerdere studies en met verschillende instrumenten wat betreft een significante verlaging van risicofactoren gedurende de behandeling.⁴⁷ Verder werden lagere recidivecijfers gevonden voor zedendelinquenten die behandeling hadden gehad vergeleken met zedendelinquenten met hetzelfde risicoprofiel die geen behandeling hadden gehad.⁴⁸

De conclusies uit de in dit artikel beschreven RCT's en studies sluiten grotendeels aan bij internationale bevindingen, zoals gerapporteerd in de drie internationale meta-analyses⁴⁹ en het onderstreept hoe complex het is om gedegen effectiviteitsonderzoek in de klinische forensische zorg uit te voeren.

6.2 Implicaties voor de praktijk

Uit de analyse van de aangeleverde studies kunnen – in integratie met bestaande literatuur – enkele implicaties voor de praktijk worden benoemd. Zowel door Fie-

43 Klein Tuente et al. (2020).

44 Van Alphen et al. (2019).

45 Papalia et al. (2019).

46 Zie ook Harte (2019).

47 Jankovic et al. (2021); De Vries Robbé et al. (2014).

48 Smid et al. (2016).

49 Abbing et al. (2023); McIntosh et al. (2021); Papalia et al. (2019).

Vivienne de Vogel e.a.

lenbach et al.⁵⁰ als door Bloemers en De Vogel⁵¹ werd gerapporteerd dat patiënten moeite leken te hebben om de behandelprogramma's te volgen (het lukte niet binnen de tijd, ze hadden moeite om huiswerkopdrachten te maken). Gezien de complexiteit van de problemen van de forensische doelgroep is dit niet verwonderlijk en blijft het dus essentieel om deze patiënten te stimuleren en hulp en ondersteuning te bieden om de behandeling te volgen zoals gepland. Ook in het onderzoek van Klein Tuente et al.⁵² werd het belang benadrukt van ondersteuning van patiënten bij het oefenen met nieuwe vaardigheden in de dagelijkse praktijk. Dit vereist meer dan enkel de therapieessies van de behandelvorm, maar voortzetting van oefenen op bijvoorbeeld de leefgroepen. Ook in de meta-analyse van Papalia et al.⁵³ werd dit als mogelijke succesfactor benoemd. Dit betekent dat staf in bredere zin getraind moet worden en dat er aandacht moet zijn voor goede onderlinge communicatie. Verder is duidelijk dat de meeste behandelprogramma's nog beter kunnen worden aangepast aan de forensische context of kunnen worden gemoderniseerd. Multimodale intensieve behandeling lijkt uiteindelijk het meest effectief; een combinatie van bijvoorbeeld cognitief gedragstherapeutische therapie, psycho-educatie en creatieve of bewegings-behandelvormen.

6.3 *Beperkingen huidige compilatie*

De huidige compilatie kent beperkingen. Het overzicht is mogelijk niet volledig, hoewel alle instellingen meermaals zijn bevestigd en alle hebben aangeleverd en de literatuursearch maar één extra studie opleverde. Verder bestrijkt het onderzoek een beperkte periode, namelijk van 2005 tot 2023. We hebben hiervoor gekozen, omdat de periode voor 2005 moeilijk te vergelijken is met de huidige forensische klinische zorg. Er hebben in deze periode belangrijke ontwikkelingen plaatsgevonden, zowel inhoudelijk in de vorm van nieuwe, meer geprotocolleerde behandelingen, de verplichtstelling van gestructureerde risicotaxatie, maar ook veranderingen in populatie, en wijzigingen in de organisatie van de forensische zorg. Tot slot moet worden vermeld dat deze compilatie puur beschrijvend is. Een gedegen meta-analyse wordt als niet mogelijk gezien vanwege de grote verschillen in type onderzoek, populatie, type interventie, gebruikte uitkomstmaten en follow-upperiode.

6.4 *Suggesties voor vervolgonderzoek*

Uit de huidige compilatie kunnen diverse lessen worden gehaald voor vervolgonderzoek, die om meer concrete uitwerking vragen, zowel qua inhoud, onderzoeksdesign als proces (financiering, samenwerking, implementatie, enz.). De huidige werkgroep zal hiermee – in samenwerking met het forensisch veld, waaronder behandelaars – de komende tijd aan de slag gaan. Enkele suggesties die verder uitgedacht zullen gaan worden, zijn het opzetten van gezamenlijk onderzoek naar be-

50 Fielenbach et al. (2018).

51 Bloemers & De Vogel (2020).

52 Klein Tuente et al. (2020).

53 Papalia et al. (2019).

handeling in de klinische forensische zorg, zowel naar TAU als naar specifieke behandelprogramma's.

Qua onderzoeksdesign wordt gedacht aan zowel kwantitatieve als kwalitatieve studies, dan wel een combinatie daarvan. Bij kwantitatief onderzoek is het onder meer belangrijk dat patiënten over een langere periode worden gevolgd, dat er objectieve en eenduidige uitkomstmaten worden gebruikt, dat er rekening wordt gehouden met verschillende subpopulaties (o.a. gender, leeftijd, type delict, stoornis) en dat drop-out zo veel mogelijk moet worden voorkomen. Samenwerking met verschillende instellingen is hierbij essentieel. Relevante vragen zijn niet alleen of behandeling effectief is, maar ook voor wie precies, wanneer en onder welke omstandigheden. Met behulp van kwalitatief onderzoek is het mogelijk om ervaringen van patiënten en behandelaars meer diepgaand te onderzoeken. Soms zijn het relatief kleine dingen die een groot verschil maken, zoals een bepaalde therapeut of een specifieke gebeurtenis die een patiënt inzicht gaf. Dergelijke zaken zijn moeilijk te meten in kwantitatief onderzoek. Meer in het algemeen is er recentelijk een tendens naar meer kwalitatief onderzoek. Zo startten het Nederlands Jeugdinstituut (NJI) en andere kennisinstituten in het Samenwerkingsverband Erkenning van Interventies in augustus 2023 een pilot naar (het combineren van kwantitatief en) kwalitatief onderzoek om beter inzicht te krijgen in de mogelijkheden om met dit type onderzoek behandel-effectiviteit aan te tonen.⁵⁴

Het is voorts van belang niet alleen te kijken naar het behandelprogramma, maar ook naar kernmerken van de patiënt (bijv. motivatie, diagnose, gender), de werkrelatie tussen professionals en patiënten en naar de context, hoe behandeling in het algemeen wordt aangeboden. Tot slot zou het interessant zijn om niet alleen te kijken naar de forensische klinische zorg, maar naar behandel-effectiviteit gedurende het gehele forensische traject, dus bijvoorbeeld ook bij ambulante behandeling of bij doorstroom naar andere zorgaanbieders, of toezichthouders zoals de reclasering.

7 Conclusie

Recidivecijfers na ontslag zijn relatief gunstig en recidive tijdens verlof is zeldzaam. Ook hebben meerdere onderzoeken een significante verlaging van risicofactoren gevonden en verhoging van beschermende factoren. Wat precies werkt, voor wie en wanneer is echter nog onduidelijk, en onderzoek naar effectiviteit van behandeling is complex om uit te voeren. In zes RCT's gepubliceerd tussen 2005 en 2023 werden voor verschillende behandelingen kleine of kortetermijneffecten gevonden, maar alle RCT's hadden ernstige tekortkomingen, waardoor nog geen definitieve conclusie over behandel-effectiviteit kan worden getrokken. Opvallend is dat ook voor TAU effecten werden gevonden. Voor toekomstig onderzoek is het van belang TAU te onderzoeken en naast RCT's ook andere vormen van onderzoek uit te voeren, waaronder meer kwalitatief onderzoek. Aandacht voor de context,

54 Zie www.nji.nl/nieuws/kwalitatief-onderzoek-bij-beoordelen-interventies.

Vivienne de Vogel e.a.

kenmerken van de patiënt en de behandelrelatie is hierbij essentieel, net als samenwerking tussen onderzoekers van verschillende instellingen.

Literatuur

- Abbing, A., Haeyen, S., Nyapati, S., Verboon, P., & Hooren, S.V. (2023). Effectiveness and mechanisms of the arts therapies in forensic care. A systematic review, narrative synthesis, and meta analysis. *Frontiers in Psychiatry*, 14. doi.org/10.3389/fpsy.2023.1128252
- Alphen, R. van, Stams, G.J.J.M., & Hakvoort, L. (2019). Musical Attention Control Training for psychotic psychiatric patients: An experimental pilot study in a forensic psychiatric hospital. *Frontiers in Neuroscience*, 13, 570. doi.org/10.3389/fnins.2019.00570
- Bernstein, D., Keulen-de Vos, M., Clerck, M., Vogel, V. de, Kersten, G.C.M. Lancel, M., Jonkers, P.P., Bogaerts, S., Slaats, M., Broers, N.J., Deenen, T.A.M., & Arntz, A. (2023). Schema therapy for violent PD offenders: A randomized clinical trial. *Psychological Medicine*, 53(1), 88-102. doi.org/10.1017/S0033291721001161
- Bloemers, E., & Vogel, V. de (2020). COMET voor zelfbeeld en paranoia bij paranoïde patiënten. Een kleinschalige RCT binnen de forensische psychiatrie. *Gedragstherapie*, 53(2), 82-102.
- Bongers, I.L., Buitenweg, D.C., & Nieuwenhuizen, Ch. van (2021). *Welke longitudinale onderzoeksvragen zijn relevant voor het forensisch psychiatrisch veld: Een klassieke Delphi-studie*. KFZ. <https://kfz.nl/projecten/precall-longitudinaal-onderzoek-call-2019-97>
- Bulten, E., Vogel, V. de, & Kordelaar, W. van (2022). Onttrekkingen en recidive tijdens tbs-verlof: De wet van de kleine aantallen. *Tijdschrift voor Psychiatrie*, 64, 657-662.
- Fielenbach, S., Donkers, F.C., Spreen, M., & Bogaerts & S. (2018). Effects of a theta/sensorimotor rhythm neurofeedback training protocol on measures of impulsivity, drug craving, and substance abuse in forensic psychiatric patients with substance abuse: Randomized controlled trial. *JMIR Mental Health*, 5(4). doi.org/10.2196/10845
- Gerstman, B.B. (2023). There is no single gold standard study design (RCTs are not the gold standard). *Expert Opinion on Drug Safety*, 22, 267-270. doi.org/10.1080/14740338.2023.2203488
- Harte, J. (2019). *Zo werkt het. Over hoe onderzoek bijdraagt aan betere interventies* (oratie). Boom criminologie.
- Hong, Q.N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman F., Cargo, M., Dagenais, P., Gagnon, M-P, Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018. Registration of Copyright (#1148552)*. Canadian Intellectual Property Office, Industry Canada.
- Janković, M., Boxtel, G.V., Masthoff, E., Caluwé, E.D., & Bogaerts, S. (2021). The long-term changes in dynamic risk and protective factors over time in a nationwide sample of Dutch forensic psychiatric patients. *Frontiers in Psychiatry*, 12, 737846. doi.org/10.3389/fpsy.2021.737846
- Klein Tuente, S., Bogaerts, S., Bulten, E., Keulen-de Vos, M.E., Vos, M., Bokern, H., van IJzendoorn, S., Geraets, C.N.W., & Veling, W. (2020). Virtual Reality aggression prevention therapy (VRAPT) versus waiting list control for forensic psychiatric inpatients: A multicenter randomized controlled trial. *Journal of Clinical Medicine*, 9, 2258. doi.org/10.3390/jcm9072258 / <https://www.mdpi.com/2077-0383/9/7/2258>

- McIntosh, L.G., Janes, S., O'Rourke, S., & Thomson, L.D. (2021). Effectiveness of psychological and psychosocial interventions for forensic mental health inpatients: A meta-analysis. *Aggression and Violent Behavior, 58*, 101551. doi.org/10.1016/j.avb.2021.101551
- Nagtegaal, M.H., Goethals, K., & Meynen, G. (2016). De tbs-maatregel: Kosten en baten in perspectief. *Tijdschrift voor Psychiatrie, 58*(10), 739-745.
- Neijmeijer, L.J., Korzilius, H.P.L.M., Kroon, H., Nijman, H.L.I., & Didden, R. (2019). Flexible assertive community treatment for individuals with a mild intellectual disability or borderline intellectual functioning: results of a longitudinal study in the Netherlands. *Journal of Intellectual Disability Research, 63*(8), 1015-1022. doi.org/10.1111/jir.12619
- Papalia, N., Spivak, B., Daffern, M., & Ogloff, J.R. (2019). A meta-analytic review of the efficacy of psychological treatments for violent offenders in correctional and forensic mental health settings. *Clinical Psychology: Science and Practice, 26*(2). doi.org/10.1111/cpsp.12282
- Schuringa, E., Spreen, M., & Bogaerts, S. (2019). Inpatient violence in forensic psychiatry: Does change in dynamic risk indicators of the IFTE help predict short term inpatient violence? *International Journal of Law and Psychiatry, 66*, 101448. doi.org/10.1016/j.ijlp.2019.05.002
- Smid, W.J., Kamphuis, J.H., Wever, E.C., & Beek, D.J. van (2016). A quasi-experimental evaluation of high-intensity inpatient sex offender treatment in the Netherlands. *Sexual Abuse, 28*(5), 469-485. doi.org/10.1177/1079063214535817
- Vogel, V. de, Bruijn, M. de, Klein Haneveld, E., Leong, J., Robbe, E., Steunenbergh, D., Verhees, M., & Vissers, A. (2023). *Behandelen van vrouwen: Een vak apart? Ontwikkeling van een handreiking voor het behandelen van vrouwen in de klinische forensische zorg*. Utrecht: Kwaliteit Forensische Zorg (KFZ).
- Vries Robbé, M. de, Vogel, V. de, Douglas, K.S., & Nijman, H.L.I. (2014). Changes in dynamic risk and protective factors for violence during inpatient forensic psychiatric treatment: Predicting reductions in postdischarge community recidivism. *Law and Human Behavior, 39*, 53-61. doi.org/10.1037/lhb0000089
- Wijma, H., & Bulten, E. (2021). How to assess recovery in a forensic population: The psychometric qualities of the Recovery Assessment Scale 24 (RAS-24). *Journal of Recovery in Mental Health, 4*(2), 4-15. https://jps.library.utoronto.ca/index.php/rmh/article/view/36975
- Wolf, M.J.F. van der, Reef, J., & Wams, A.C. (2020). *Wie zijn geschiedenis niet kent...: Een overzichtelijke tijdlijn van de stelselwijzigingen in de forensische zorg sinds 1988*. Instituut voor Strafrecht en Criminologie, Universiteit Leiden.
- Zwets, A.J., Hornsveld, R.H.J., Muris, P., Kanters, T., Langstraat, E., & Marle, H.J.C. van (2016). Psychomotor Therapy as an additive intervention for violent forensic psychiatric inpatients: A pilot study. *International Journal of Forensic Mental Health, 15*(3), 222-234. doi.org/10.1080/14999013.2016.1152613