

Gedragsveranderingen tijdens tbs-behandeling: een multicenteronderzoek

E. DE JONGE, H.L.I. NIJMAN, S.M.M. LAMMERS

ACHTERGROND In de forensische psychiatrie is het zo betrouwbaar mogelijk inschatten van de recidiverisico's van patiënten cruciaal. Het risicotaxatie-instrument HKT-30 brengt zowel statische als dynamische risicofactoren in kaart. De dynamische risicofactoren zouden door behandeling ten goede moeten veranderen.

DOEL Onderzoeken of dynamische factoren veranderen tijdens tbs-behandeling.

METHODE De HKT-30 werd in totaal 984 maal gescoord bij tbs-patiënten in 3 verschillende forensisch psychiatrische centra, medio 2003-november 2006. Voor 513 patiënten gebeurde dat ten minste 1 maal, voor 313 2 maal of meer en voor 158 in 3 opeenvolgende jaren. Opdeling van de onderzoeksgroep in 'nieuwe', 'oude' en 'overgeplaatste' patiënten maakte het mogelijk te onderzoeken in welke fase van de behandeling de meeste verandering in scores plaatsvond.

RESULTATEN Een groot deel van de scores op de dynamische risicofactoren, en daarmee ook de totaalscore, namen statistisch significant af, maar de absolute verschillen waren klein. De drie groepen patiënten verschilden niet veel van elkaar in mate van verandering.

CONCLUSIE De scores op de in theorie als veranderbaar veronderstelde risicofactoren verminderden naarmate de behandeling vorderde. Daarmee staat echter nog niet vast of dit ook daadwerkelijk een vermindering van het recidiverisico betekent.

[TIJDSCHRIFT VOOR PSYCHIATRIE 51(2009)4, 205-215]

TREFWOORDEN dynamische risicofactoren, forensische psychiatrie, HKT-30, risicotaxatie, tbs

In de forensische psychiatrie is het zo betrouwbaar mogelijk inschatten van recidiverisico's van patiënten cruciaal. Risicotaxatie vormt de basis voor belangrijke beslissingen zoals het verlenen van verlop aan tbs-patiënten en het verlengen van de maatregel terbeschikkingstelling (tbs) (Brand & Diks 2001). In het verleden voer men op de klinische inschatting van het recidiverisico. Echter, in de afgelopen decennia wordt er wegens de onbetrouwbaarheid van deze inschatting steeds meer gebruik gemaakt van gestructureerde methoden van risicotaxatie (Lammers 2007; Van Nieuwenhuizen & Philipse 2002; De Vogel 2005; raadpleeg

deze artikelen ook voor uitgebreidere achtergronden over risicotaxatie).

Bij sommige van deze gestructureerde methoden gebruikt men uitsluitend statische, niet (ten positieve) veranderbare factoren om tot een eendoordeel te komen. De praktische waarde voor behandelaars bij behandeling en begeleiding is daarmee beperkt (De Vogel e.a. 2003). Onderzoekers van een aantal tbs-klinieken ontwikkelden het risicotaxatie-instrument HKT-30 (van 'historisch-klinisch-toekomst'; Werkgroep Risicotaxatie Forensische Psychiatrie 2002). Hun doelstelling was om een instrument te ontwikkelen dat beter

aansluit bij de situatie in de Nederlandse forensische psychiatrie dan de reeds bestaande instrumenten.

De HKT-30 bestaat, behalve uit 11 statische factoren, uit 19 dynamische variabelen; deze zijn potentieel veranderbaar door behandeling en begeleiding (zie tabel 1). De dynamische, maar ook de historische (statische) schaal van de HKT-30 lijkt een redelijke voorspellende waarde (predictieve validiteit) te hebben voor gewelddadige recidive (Hildebrand e.a. 2005, 2006). In het quasi-prospectieve onderzoek van Hildebrand e.a. (2005) bleek de totaalscore van de dertig items de beste voorspeller voor gewelddadige recidive (*area under the curve of AUC = 0,72*). De AUC is de uitkomstwaarde bij analyse met *receiver-operating characteristics* en geeft kortgezegd de voorspellende waarde. Een waarde van 0,50 staat gelijk aan toeval en waarden van $\geq 0,75$ worden gezien als goed (Hildebrand e.a. 2005). In een onderzoek naar ontstekingen en ontvluchtelingen tijdens verlof en daarmee samenhangende gewelddadige recidive vonden Hildebrand e.a. (2006) een veelbelovende voorspellende waarde voor een combinatie van negen K-items ($AUC = 0,83$), maar een replicatiestudie is nodig om de waarde van deze bevinding nader te onderzoeken.

De dynamisch geachte items van de HKT-30 zijn dat uiteraard alleen echt als ze ook daadwerkelijk kunnen veranderen in de loop van de tijd. Aangezien de HKT-30 nu sinds zo'n 4 jaar in een aantal tbs-klinieken jaarlijks wordt gescoord, is het mogelijk verschuivingen in scores in de loop van de behandeling te onderzoeken. Een eerder onderzoek met de *Rehabilitation Evaluation Hall and Baker* (REHAB; Baker & Hall 1988; Van der Gaag & Wilken 1994) wees erop dat het gedrag van patiënten vooral in de eerste periode van de tbs-behandeling verbeterd (Nijman e.a. 2004).

Daarom verrichtten wij een onderzoek waarbij wij onderscheid maakten tussen patiënten voor wie kort na binnenkomst de eerste HKT-30 werd gescoord en patiënten die al langer in de kliniek verbleven op het moment dat men begon met het gebruik van de HKT-30. De eerste groep ver-

deelden wij vervolgens verder in twee subgroepen: enerzijds de echt 'nieuwe' patiënten, dit wil zeggen degenen die onlangs in de tbs werden opgenomen, en anderzijds de overgeplaatste patiënten. Deze laatstgenoemden, in tbs-jargon 'herselectanten' genoemd, zijn na een vastgelopen behandeling in een andere tbs-kliniek overgeplaatst naar de huidige kliniek. Zij maken dus een nieuwe start, maar het zijn geen 'nieuwe' tbs-patiënten.

De hypothesen van dit onderzoek luiden als volgt:

(a) De somscores op de klinische subschaal en de toekomstsubschaal nemen voor de totale groep patiënten af over het verloop van de tijd. Hierdoor neemt ook de totaalscore af.

(b) Bij de eerste meting (T_1) scoren de oude patiënten, die immers al langer in behandeling zijn, lager dan de nieuwe patiënten.

(c) Bij de eerste meting scoren overgeplaatste patiënten (in het vervolg 'overgeplaatsten' genoemd) hoger dan de andere patiënten, omdat van hen al is gebleken dat zij moeilijk behandelbaar zijn.

(d) Om dezelfde reden nemen de scores van overgeplaatsten minder snel af over het verloop van de tijd.

(e) De afname van scores is het sterkst voor de nieuwe patiënten.

METHODE

Onderzoeksinstrument De items van de HKT-30 zijn weergegeven in tabel 1. De scores op de items lopen van 0 tot 4, waarbij een score 0 het laagste risiconiveau weergeeft. De historische items betreffen gedrag en kenmerken uit het verleden van de patiënt. De klinische variabelen en toekomstvariabelen dienen, op één item na, gescoord te worden over het afgelopen jaar. De interne consistentie van de drie schalen is voldoende tot goed, waarbij de historische schaal wat lijkt achter te blijven bij de andere twee schalen (Werkgroep Risicotaxatie Forensische Psychiatrie 2002; Hildebrand e.a. 2005). De interbeoordelaarsbetrouwbaarheid van de subschalen en van de to-

taalscore is, volgens een aantal onderzoeken uitgevoerd met oude dossiers, voldoende tot goed, maar laat voor enkele individuele items wat te wensen over (Canton e.a. 2004; Hildebrand e.a. 2005; 2006; Werkgroep Risicotaxatie Forensische Psychiatrie 2002). In de praktijk van de risicotaxatie blijkt er echter voor bijna alle individuele items een goede tot zeer goede interbeoordelaarsbetrouwbaarheid te zijn (Groot Rouwen 2006 in Lammers 2007). Zoals eerder vermeld, lijkt de HKT-30 een redelijk goede predictieve validiteit te hebben waar het toekomstige (gewelddadige) recidive betreft.

Dataverzameling Voor alle (mannelijke) patiënten die tussen medio 2003 en november 2006 verbleven in de forensisch psychiatrische centra de Oostvaarderskliniek, de Kijvelanden en Oldenkotte werd een HKT-30-beoordeling uitgevoerd. Patiënten die na 2004 werden opgenomen, werden zo snel mogelijk na binnenkomst beoordeeld. Beoordelingen en herbeoordelingen vonden in de meeste gevallen jaarlijks plaats: bij het aanvragen van een (andere) verlofstatus, bij verlengingsadviezen of voor de jaarlijkse evaluatie van het verlof.

Analyses Voor de toetsing van de verschillen in scores op T1 werd gebruikgemaakt van variantieanalyse (eenzijdige ANOVA). De verschillen in scores over het verloop van de tijd werden geanalyseerd met lineaire gemengde modellen (LMM-methode). Het voordeel van deze methode boven het *repeated measures general linear model* (GLM) is dat in de LMM ook personen met minder dan 3 metingen kunnen worden meegenomen. Bij de GLM zouden dezen geheel buiten de analyse gelaten worden. Bij missende waarden, zoals hier het geval was, worden gemiddelden en vrijheidsgraden (df) geschat. Gemiddelden zullen daarom anders zijn dan de werkelijke waarden en vrijheidsgraden zijn niet altijd hele getallen. Voor de toetsing van het beloop voor de gehele groep patiënten werd het hoofdeffect van de metingen gebruikt. Voor de toetsing van

de verschillen in het beloop tussen de 3 groepen werd ook het hoofdeffect van de groepen meegenomen, evenals de interactie tussen de metingen en de groepen. Bij het gebruik van post-hoc-toetsen werd een bonferronicorrectie uitgevoerd.

RESULTATEN

Kenmerken van de patiënten In totaal werden de gegevens van 984 metingen verzameld. Voor 513 patiënten werd ten minste 1 maal een HKT-30 ingevuld en voor 313 gebeurde dat 2 maal of meer; voor 158 patiënten werd 3 maal een HKT-30 beoordeeld. De gemiddelde leeftijd van de patiënten ($n = 513$) was op het moment van de eerste HKT-30-beoordeling (T1) 37,6 jaar (SD: 9,4). Het gemiddelde IQ bedroeg 96,7 (SD: 14,8). Bij 68,3% van deze patiënten was sprake van midde-lenafhankelijkheid, bij 29,9% van een psychotische stoornis en bij 60,6% van een (cluster B-) persoonlijkheidsstoornis. Bij het merendeel van de patiënten die niet bekend waren met één van deze stoornissen (10,2%), was er sprake van een autistische stoornis of pedofilie.

De maatregel terbeschikkingstelling werd in de onderzoeksgroep opgelegd voor (poging tot) moord of doodslag (40,7%), verkrachting of aanranding van een slachtoffer van 16 jaar of ouder (14,6%) of jonger dan 16 jaar (13,3%), mishandeling of geweldpleging (10,3%), een vermogensdelict (al dan niet met geweld; 11,7%) of brandstichting (8,4%). Het percentage patiënten dat een eerdere tbs-maatregel had gehad, bedroeg 7,5. Daarnaast had 30% van de patiënten een opname in een andere tbs-kliniek achter de rug vóór de opname in de huidige kliniek (de groep overgeplaatste patiënten).

Op T1 verbleven de patiënten gemiddeld 50,3 (SD: 44,8) maanden in een tbs-kliniek, waarvan 30,9 maanden (SD: 34,0) in de huidige kliniek. Tussen de 1ste en de 2de HKT-30-beoordeling (T2) zat een periode van gemiddeld 13,6 maanden (SD: 5,8) en tussen de 2de en de 3de beoordeling (T3) een periode van 11,0 maanden (SD: 3,1).

Voor 128 patiënten werd de HKT-30-beoorde-

ling uitgevoerd binnen 1 jaar nadat zij in zowel de tbs als de kliniek werden opgenomen (de zogenaamde 'nieuwe patiënten'). Voor 36 van hen waren ook een 2de en een 3de meting verricht. Daarnaast waren er 272 patiënten voor wie de 1ste beoordeling pas ná 1 jaar opname plaatsvond (de zogenaamde 'oude patiënten'). Deze groep verbleef op T1 gemiddeld al 4,9 jaar (SD: 3,4) in de tbs. Voor 99 van hen werden ook een 2de en een 3de meting uitgevoerd. Ten slotte waren er nog 76 patiënten voor wie T1 viel op het moment dat zij al langer dan 1 jaar in de tbs-kliniek verbleven, maar minder dan 1 jaar in de huidige kliniek (de 'overgeplaatste patiënten'). Voor 23 patiënten uit deze groep werden ook een 2de en een 3de beoordeling uitgevoerd. Deze groep verbleef op T1 gemiddeld 5,9 jaar (SD: 2,8) in de tbs.

De verschillen in persoonskenmerken tussen patiënten uit deze drie groepen werden getoetst met behulp van variantieanalyses. Er bleken alleen verschillen te zijn in IQ-score ($F[2; 386] = 3,56; p = 0,030$). Post hoc bleken de nieuwe patiënten lagere scores te hebben dan de overgeplaatste patiënten ($p = 0,030$). Mogelijk hing dit verschil samen met het gebruik van andere testinstrumenten; vooral het vroegere gebruik van de Wechsler Adult Intelligence Scale (WAIS-II) en tegenwoordig van de WAIS-III.

Hypothese 1 De somscores van de klinische subschaal en de toekomstsubschaal, evenals de HKT-30-totaalscore, namen voor de gehele groep af naarmate de tijd verstreek (resp. $F[2; 407,65] = 23,75; p \leq 0,001$; $F[2; 450,98] = 22,64; p \leq 0,001$ en $F[2; 381,62] = 20,90; p \leq 0,001$). De score op de historische schaal nam echter toe met het verstrijken van de tijd ($F[2; 475,71] = 8,63; p \leq 0,001$).

Zoals eerder vermeld, worden de gemiddelde waarden bij gebruik van de LMM-methode geschat. Deze gemiddelde waarden zijn vermeld in tabel 2. Aan de hand hiervan, en met de overige waarden die in de tabel vermeld worden, berekenen wij de effectgroottes (Cohens d , zie ook tabel 2). Deze varieerden van zeer klein voor de historische schaal tot klein-gemiddeld voor de toekomstschaal.

Wij zagen significante verbeteringen op 8 klinische items en op alle toekomstitems. Op de historische schaal was een verandering te zien voor 1 item; er was namelijk een toename in het aantal schendingen van voorwaarden omtrent behandeling en toezicht (Ho2). De klinische items die afnamen, waren: gebrek aan probleeminzicht (Ko1), impulsiviteit (Ko4), gebrek aan empathie (Ko5), vijandigheid (Ko6), gebrek aan sociale en relationele vaardigheden (Ko7), attitude ten opzichte van de behandeling (K10), gebrek aan verantwoordelijkheid voor het delict (K11) en gebrek aan copingvaardigheden (K13). Van de toekomstitems namen de gemiddelde scores af op: gebrek aan overeenstemming over voorwaarden (To1), materiële indicatoren (To2), ontbrekende dagbesteding (To3), gebrek aan vaardigheden (To4), sociale steun en netwerk (To5) en stresserende omstandigheden (To6). De vrijheidsgraden, F-waarden en het significantieniveau zijn weergegeven in tabel 1.

Hypothesen 2 en 3 De totaalscore van de HKT-30 bleek op T1 te verschillen tussen de drie groepen ($F[2; 899] = 4,35, p = 0,013$). Post hoc bleek echter niet dat de oude patiënten lagere scores hadden dan de nieuwe patiënten (hypothese 2). Overgeplaatsten scoorden hoger dan de groep oude patiënten ($p = 0,011$), maar verschilden niet met de nieuwe patiënten ($p = 0,058$). Hypothese 3 kon hiermee slechts gedeeltelijk bevestigd worden. Echter, op de toekomstsubschaal bleken er wel verschillen tussen de groepen te zijn ($F[2; 900] = 42,35, p \leq 0,001$). De oude patiënten scoorden hier beter dan de overgeplaatste ($p \leq 0,001$) en de nieuwe patiënten ($p \leq 0,001$), die onderling niet verschilden. Ook op de historische schaal verschilden de groepen ($F[2; 941] = 18,65; p \leq 0,001$). Hier hadden de nieuwe patiënten lagere scores dan de oude ($p \leq 0,001$) en de overgeplaatste patiënten ($p \leq 0,001$). De klinische subschaal verschilde niet voor de verschillende groepen ($F[2; 935] = 1,44; p = 0,24$).

TABEL 1 Verandering in scores op itemniveau op de HKT-30 voor de totale groep over een periode van drie metingen, getoetst met lineair gemengd model ($n \geq 498$), bij tbs-gestelden in 3 verschillende klinieken, medio 2003-november 2006

Code	Item	df	F-waarde	p
Historische factoren				
H01	Justitiële voorgeschiedenis	2; 372,95	0,01	0,996
H02	Schending van voorwaarden omtrent behandeling en toezicht	2; 388,23	8,75	0,000*
H03	Gedragsproblemen voor 12de levensjaar	2; 371,80	2,14	0,119
H04	Slachtoffer van geweld in jeugd	2; 326,47	0,89	0,411
H05	Hulpverleningsgeschiedenis	2; 379,66	1,84	0,160
H06	Arbeidsverleden	2; 367,03	1,90	0,151
H07	Middelengebruik	2; 371,54	1,12	0,327
H08	Psychotische stoornissen	2; 321,15	0,54	0,586
H09	Persoonlijkheidsstoornissen	2; 389,51	0,21	0,812
H10	Psychopathie	2; 370,30	1,98	0,140
H11	Seksuele deviantie	2; 381,81	1,18	0,310
Klinische factoren				
K01	Probleeminzicht	2; 403,90	30,57	0,000
K02	Psychotische symptomen	2; 375,26	1,76	0,174
K03	Middelengebruik	2; 520,67	0,19	0,829
K04	Impulsiviteit	2; 443,77	8,54	0,000
K05	Empathie	2; 438,60	18,88	0,000
K06	Vijandigheid	2; 461,21	5,29	0,005
K07	Sociale en relationele vaardigheden	2; 423,27	27,11	0,000
K08	Zelfredzaamheid	2; 427,64	1,51	0,222
K09	Acculturatieproblematiek	2; 442,79	0,61	0,544
K10	Attitude ten opzichte van behandeling	2; 408,71	3,47	0,032
K11	Verantwoordelijkheid voor het delict	2; 448,29	14,06	0,000
K12	Seksuele preoccupatie	2; 385,48	0,82	0,440
K13	Copingvaardigheden	2; 448,78	17,12	0,000
Toekomstige factoren				
To1	Overeenstemming over voorwaarden	2; 483,18	12,02	0,000
To2	Materiële indicatoren	2; 522,28	16,26	0,000
To3	Dagbesteding	2; 502,00	8,05	0,000
To4	Vaardigheden	2; 455,30	22,53	0,000
To5	Sociale steun en netwerk	2; 414,28	5,59	0,004
To6	Stresserende omstandigheden	2; 478,86	17,61	0,000

* Negatieve richting (toename score).

TABEL 2 Geschatte gemiddelde schaalscores (standaardfout) van de HKT-30 per meetmoment, groeps grootte (n) op T1 en effectgrootte (Cohens d) voor de verschillen tussen T1 en T3

Meting	Historische schaal (0-44)	Klinische schaal (0-52)	Toekomstschaal (0-24)	HKT-30-totaalscore (0-120)
T1 (n = 513)	23,77 (0,311) (n = 513)	21,70 (0,349) (n = 511)	15,95 (0,247) (n = 499)	61,48 (0,702) (n = 498)
T2 (n = 313)	24,04 (0,316)	20,05 (0,396)	15,21 (0,283)	59,34 (0,761)
T3 (n = 158)	24,43 (0,334)	18,76 (0,501)	13,40 (0,373)	56,83 (0,905)
Cohens d (T1-T3)	0,09	0,30	0,36	0,26

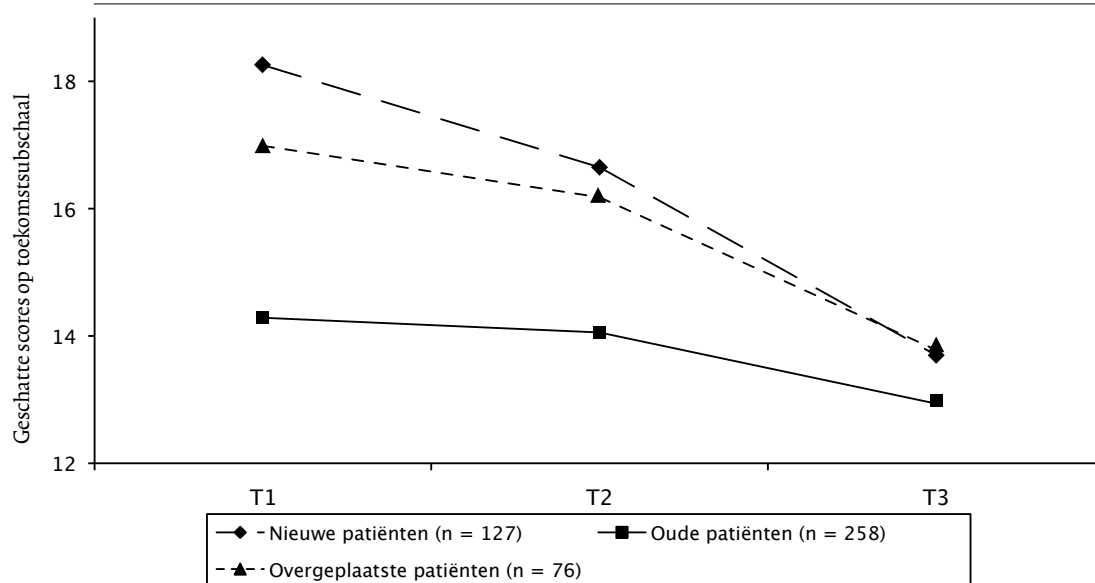
Hypothesen 4 en 5 Analyse van de schaal-scores wees uit dat de drie subgroepen alleen verschilden in veranderingen op de toekomst-subschaal ($F[4; 444,65] = 3,42; p = 0,009$). De scores van alle drie de groepen verbeterden significant op deze schaal: oude patiënten $F[2; 326,16] = 4,29; p = 0,014$, nieuwe patiënten $F[2; 323,70] = 17,36; p \leq 0,001$ en overgeplaatste patiënten $F[2; 325,45] = 5,56; p = 0,004$. Echter, terwijl de oude patiënten op T1 gunstigere scores hadden, waren er op T3 geen significante verschillen tussen de groepen meer ($F[2; 876,02] = 0,55; p = 0,576$). De oude patiënten werden gedurende de tijd dus als het ware ingehaald door de nieuwe en overgeplaatste patiënten. In figuur 1 zijn deze veranderingen grafisch weergegeven.

Op de klinische subschaal verbeterden de scores van oude en nieuwe patiënten (resp. $F[2; 286,48] = 15,87; p \leq 0,001$ en $F[2; 295,09] = 8,07; p \leq 0,001$), terwijl de scores voor overgeplaatste patiënten niet significant veranderden ($F[2; 296,45] = 0,68; p = 0,509$). Het verschil in verandering tussen de groepen was echter niet significant ($F[4; 394,29] = 0,96; p = 0,428$). Ook waren de scores van de drie groepen gemiddeld over de drie metingen gelijk ($F[2; 535,01] = 2,05; p = 0,130$).

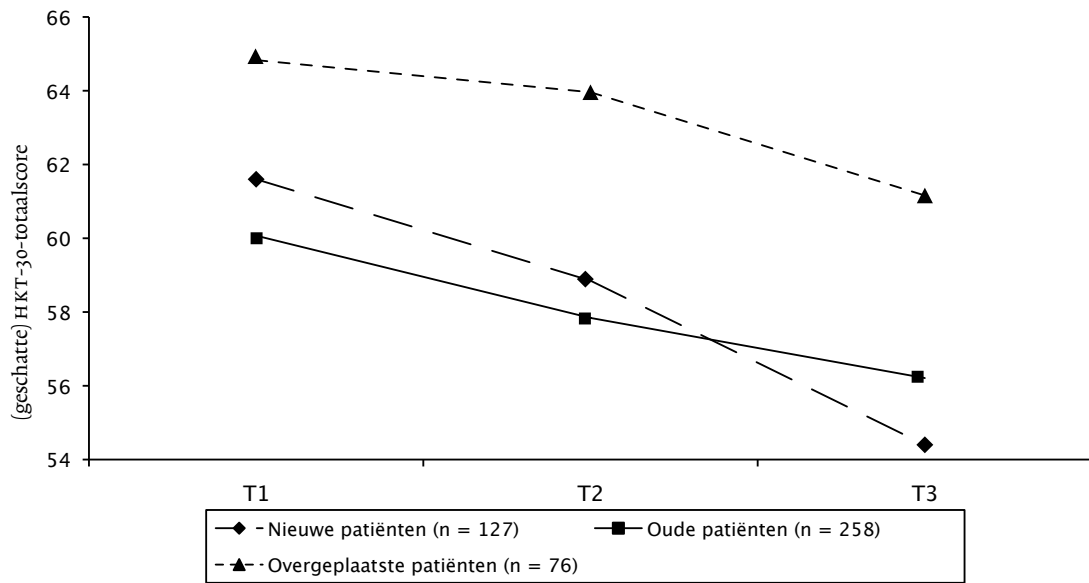
De historische subschaal gaf een vergelijkbaar beeld te zien, zij het in de tegenovergestelde richting. De oude patiënten verslechterden significant ($F[2; 293,09] = 5,46; p = 0,005$); de nieuwe en de overgeplaatste patiënten niet (resp. $F[2; 298,20] = 2,49; p = 0,084$ en $F[2; 307,29] = 2,51; p = 0,083$). Ook hier waren deze verschillen in verandering tussen de groepen niet significant ($F[4; 372,79] = 0,95; p = 0,434$). Wel hadden de oude en overgeplaatste patiënten over de 3 metingen gemiddeld hogere scores dan de nieuwe patiënten (resp. $p = 0,005$ en $p = 0,001$).

Ook voor de HKT-30-totaalscore was het beeld vergelijkbaar. Oude en nieuwe patiënten verbeterden ($F[2; 291,55] = 9,57; p \leq 0,001$ en $F[2; 279,62] = 11,24; p \leq 0,001$), maar de overgeplaatsten niet ($F[2; 282,16] = 1,87; p = 0,157$). Dit verschil in verandering was echter ook hier niet significant ($F[4; 364,67] = 1,23; p = 0,300$). Op de totaalscore hadden de overgeplaatste patiënten gemiddeld over de 3 metingen wel hogere scores dan de oude patiënten ($p = 0,032$), maar niet hoger dan de nieuwe patiënten ($p = 0,091$). Deze resultaten zijn weergegeven in figuur 2.

FIGUUR 1 Verandering in de scores op de toekomstsubschaal van de HKT-30 bij drie groepen tbs-gestelden



FIGUUR 2 Verandering in de HKT-30-totaalscores



DISCUSSIE

De doelstelling van tbs-behandeling is om patiënten zodanig te behandelen en te begeleiden dat na beëindiging van de maatregel het maatschappelijk risico dat zij vormen acceptabel is geworden. De HKT-30 is ontworpen om ook de dynamische risicofactoren die tijdens de behandeling beïnvloed kunnen worden te meten. Bevin- dingen van onder anderen Hildebrand e.a. (2005) laten zien dat de scores op deze risicofactoren significant samenhangen met latere recidive. Daarom is het relevant om te onderzoeken of deze variabelen ook werkelijk veranderen in de loop van de behandeling.

Wij hebben veranderingen in HKT-30-scores in de loop van de behandeling bestudeerd op zowel item- als (sub)schaalniveau. Daarbij maakten wij onderscheid tussen nieuwe, oude en overgeplaatste patiënten, zodat duidelijk kon worden in welke fase van de behandeling de meeste verandering plaatsvindt. Wij bespreken in het navolgende de belangrijkste resultaten.

Beperkingen van ons onderzoek

De somscores van de klinische subschaal en de toekomstsubschaal namen significant af en daarmee ook de HKT-30-totaalscore. Veel dynamische risicofactoren lijken in de praktijk dus ook veranderbaar te zijn. Ons onderzoek heeft echter een aantal methodologische beperkingen, waardoor het niet mogelijk is om zonder meer te concluderen dat deze veranderingen het effect van behandeling zijn. Zo kon er geen controlegroep of -conditie worden geïncludeerd. Daardoor zouden de verbeteringen specifiek (dit wil zeggen: niet aan de behandeling gerelateerd) kunnen zijn of zou er sprake kunnen zijn van regressie naar het gemiddelde. Verder is een belangrijke beperking van het onderzoek dat de HKT-30 werd gescoord door personen die in meer of mindere mate bij de behandeling van de desbetreffende patiënten betrokken zijn. Hierdoor kunnen allerlei verwachtingseffecten bij de observatoren een rol hebben gespeeld. Ten slotte maken ook het grote aantal toetsen en de kleine effectgroottes dat er geen verregaande conclusies uit deze studie getrokken kunnen worden.

Voor de analyse van verschillen tussen de

subgroepen is het een beperking dat de groep overgeplaatste patiënten veel kleiner was dan de andere twee groepen.

Veranderingen op itemniveau

Voor de dynamische items was er op itemniveau sprake van een significante verbetering op 8 van de 13 klinische items en op alle 6 de toekomstitems. De verbetering in empathie is het opmerkelijkst omdat er discussie bestaat over of een gebrek aan invoelend vermogen wel behandelbaar is. Volgens sommige behandelstudies verbetert empathie na behandeling, volgens andere echter niet (Hudson e.a. 2002; McGuire & Hatcher 2001). Neurobiologisch onderzoek suggereert dat gebrek aan empathie, althans bij psychopaten, een aangeboren en onveranderbaar fenomeen zou zijn (Hildebrand 2004). Wellicht kan empathie bij sommige patiënten, vooral de meer psychopathische, niet veranderen, maar bij anderen wel.

De meeste dynamisch geachte items die niet veranderden, zoals psychotische symptomen en seksuele preoccupatie, zijn slechts op een beperkt deel van de patiëntengroep van toepassing. Aangezien een substantieel deel van de populatie geen psychotische symptomen of duidelijk op de voorgrond staande seksuele preoccupaties had, kan deze in dit opzicht dus ook niet verbeteren. Mogelijke veranderingen voor deze factoren kunnen waarschijnlijk alleen aangetoond worden wanneer men ze toetst binnen de betreffende patiëntengroepen.

Ook voor de klinische factor middelengebruik is dit mogelijk het geval, aangezien klinieken alles in het werk stellen middelenmisbruik in de klinieken uit te bannen. Sowieso is de onderzochte groep in de huidige studie zowel qua diagnose als delictkenmerken en risicoaspecten erg heterogeen, zoals overigens voor de gehele tbspopulatie geldt (Van Emmerik 2001). Het is te verwachten dat voor groepen met verschillende stoornissen en delicten het verloop van HKT-30-scores verschilt. Een suggestie voor vervolgonder-

zoek luidt dan ook om te kijken naar de verschillen tussen deze groepen, bij voorkeur bij een grotere steekproef.

Naast de score op de klinische schaal en de toekomstschaal bleek ook de score op de historische schaal te veranderen: deze kan echter alleen maar toenemen met het verstrijken van de tijd. Op deze schaal is een verslechtering te zien wat betreft schendingen van voorwaarden tijdens de behandeling. Dit betekent dat (een deel van de) patiënten zich tijdens de behandeling schuldig heeft gemaakt aan schendingen van afspraken en regels zoals drugsgebruik, ordeverstoring in de kliniek, bezit van niet-toegestane goederen of onttrekking aan het toezicht van de kliniek. Dit item van de HKT-30 heeft een aanzienlijke voorspellende waarde voor latere recidive (Hildebrand e.a. 2005). Ook bleek uit ander onderzoek dat elke onttrekking aan het toezicht van de kliniek de kans op ernstige recidive in de toekomst anderhalf keer zo hoog maakt (Philipse e.a. 2006).

‘Schending van voorwaarden omtrent behandeling en toezicht’ is dus niet alleen een historisch, maar ook een dynamisch item. Dit onderstreept het belang om naast veranderingen op de klinische schaal en de toekomstschaal, ook eventuele verslechtingen op de historische schaal in de beoordelingen over verlenging en verlof te laten meewegen.

Verschillen tussen groepen

In dit onderzoek hebben wij ook gekeken naar de verschillen in veranderingen tussen de subgroepen nieuwe, oude en overgeplaatste patiënten op de verschillende subschalen en de HKT-30-totaalscore.

Historische subschaal Op de historische subschaal bleken de nieuwe patiënten gunstiger scores te hebben dan de andere twee groepen. De groep oude patiënten verslechterde, maar deze verandering was niet dusdanig groot dat er een significant verschil met de andere groepen ontstond. Deze resultaten zijn wellicht te verklaren

door het feit dat de oude en de overgeplaatste patiënten een problematischer verleden hebben. De nieuwe patiënten hebben voor een deel relatief lichte problematiek en zullen sneller uitstromen. Deze laatste groep zal dan lagere scores hebben op de historische subschaal.

Klinische subschaal Voor de klinische subschaal werd op T1 geen verschil gevonden tussen de drie groepen. Alleen de oude en de nieuwe patiënten verbeterden gedurende de tijd, maar de mate van verandering week niet significant af van die van de overgeplaatste patiënten.

Deze resultaten zijn in tegenspraak met de van te voren geformuleerde hypothesen. Wij verwachtten dat de oude patiënten de laagste scores zouden hebben en de overgeplaatste patiënten de hoogste scores. Ook verwachtten wij dat de nieuwe patiënten sneller zouden verbeteren dan de andere patiënten.

Dat deze hypothesen niet werden ondersteund, kan verschillende oorzaken hebben. Een eerste mogelijkheid is dat er sprake is van een validiteitsprobleem. Wellicht trad er bij de scoring van de HKT-30, en vooral bij de klinische schaal, een bias op. Bij de klinische items is, zoals de naam al aangeeft, een klinische inschatting nodig. Beoordelaars zouden de neiging kunnen hebben bij een eerste beoordeling het verleden van de patiënt, over een langere periode dan het voorgeschreven jaar, mee te wegen. Veranderingen die zich recent hebben voorgedaan, zouden dan pas bij de herbeoordeling in de scores tot uitdrukking komen. Een andere mogelijkheid is dat de oude patiënten pas na langere tijd op het niveau van de nieuwe patiënten zitten. Dit kan enerzijds veroorzaakt worden doordat de oude patiënten meer problemen hebben en moeilijker behandelbaar zijn, maar anderzijds ook door veranderende behandelmethoden.

Toekomstschaal De toekomstschaal was de enige schaal waarop de drie groepen aanzienlijk van elkaar verschilden. In overeenstemming met de hypothese bleken de oude patiënten op T1 la-

gere scores te hebben dan de andere twee groepen. Deze laatste twee groepen zijn pas kort in de kliniek. Hierdoor zijn er dus nog weinig voorbereidingen voor eventuele resocialisatie getroffen. De nieuwe en de overgeplaatste patiënten bleken in de loop van de tijd echter sneller te verbeteren, waardoor zij na 2 jaar op hetzelfde niveau zaten als de oude patiënten.

HKT-30-totaalscore Op de HKT-30-totaalscore bleken de overgeplaatste patiënten hoger te scoren dan de oude patiënten. Alleen oude en nieuwe patiënten verbeterden significant in de loop van de tijd, maar deze verbeteringen verschilden niet significant met het verloop van scores van de overgeplaatste patiënten. De groep overgeplaatste patiënten was echter veel kleiner dan de andere twee groepen, waardoor het onderscheidend vermogen voor deze groep lager was. Dit betekent dat voor deze groep een groter absoluut verschil nodig is alvorens het een statistisch significante waarde bereikt dan voor de andere groepen. De resultaten dienen daarom ook met voorzichtigheid te worden geïnterpreteerd.


Ons onderzoek betreft dan ook een eerste verkenning van de veranderbaarheid van risicofactoren, gedurende een relatief korte behandelperiode van ruim 2 jaar. Niettemin vormen de gevonden verschillen wel een indicatie dat er veranderingen tijdens de behandeling optreden. In absolute zin zijn deze niet heel groot, maar dit is ook begrijpelijk. Gedragsveranderingen komen doorgaans langzaam tot stand, zeker bij deze moeilijke groep. Niet voor niets duurt de gemiddelde tbs-behandeling meer dan 7 jaar (Van der Heide & Eggen 2007). Het zou dan ook interessant zijn om de rest van de behandeling van deze patiënten te volgen.

De vraag 'in welke fase van de behandeling de meeste verandering in risicofactoren plaatsvindt' is, gezien de genoemde beperkingen, moeilijk te beantwoorden. Toch lijkt het er in grote lijnen op dat veranderingen gedurende de gehele behandeling plaats kunnen vinden. Daarnaast kunnen wij, met de nodige voorzichtigheid, stellen dat de overgeplaatste patiënten het wat slechter lijken te

doen. Het verdient aanbeveling om deze groep in de toekomst nader te bestuderen, zeker gezien recente bevindingen waaruit bleek dat deze groep veel overeenkomsten vertoont met long-staypatiënten (Van der Kraan e.a. in druk).

BESLUIT

Hoewel de in ons onderzoek gevonden afname in dynamische risicofactoren van de HKT-30 hoopgevend is, betekent dit nog niet dat de tbs-behandeling succesvol is of dat de mate van afname ook samengaat met reductie van recidive. Zoals ook Lammers (2007) aangeeft, zijn de aange- toonde veranderingen pas écht waardevol wanneer deze blijken samen te hangen met vermindering van delictgevaar. Behandelplannen en risicobeheersing zouden dan directer gekoppeld kunnen worden aan dynamische risicofactoren. De huidige jaarlijkse dataverzameling met de HKT-30 maakt in de toekomst dergelijk onderzoek mogelijk. Daarbij zou men ook een vergelijkend onderzoek met gedetineerden binnen het gevangeniswezen als controlegroep kunnen verrichten, ook al zijn de populaties niet identiek.

 E.F.J.M. Brand, T.P.C. Lucker, A.J. Zwets en M.R.D. Harbers leverden een bijdrage aan de totstandkoming van dit artikel.

LITERATUUR

- Baker, R.D., & Hall, J.N. (1988). REHAB: a new assessment instrument for chronic psychiatric patients. *Schizophrenia Bulletin*, 14, 97-111.
- Brand, E.F.J.M., & Diks, G.J.M. (2001). Richtlijnen voor risicotaxatie in de forensische diagnostiek: theorie en praktijk. *Tijdschrift voor Psychiatrie*, 43, 693-704.
- Canton, W.J., van der Veer, T.S., van Panhuis, P.J.A., e.a. (2004). De betrouwbaarheid van risicotaxatie in de pro Justitia rapportage. Een onderzoek met behulp van de HKT-30. *Tijdschrift voor Psychiatrie*, 46, 537-542.
- Emmerik, J.L. van (2001). *De terbeschikkingstelling in maat en getal. Een beschrijving van de tbs-populatie in de periode 1995-2000*. Den Haag: Dienst Justitiële Inrichtingen.
- Gaag, M. van der, & Wilken, J. (1994). REHAB. Een maat van psychiatrische invaliditeit. Lisse: Swets & Zeitlinger.
- Heide, W. van der, & Eggen, A.T.J. (2007). *Criminaliteit en rechtshandhaving 2006. Ontwikkelingen en samenhangen*. Den Haag: WODC.
- Hildebrand, M. (2004). *Psychopathy in the treatment of forensic psychiatric patients. Assessment, prevalence, predictive validity and clinical implications*. [Proefschrift]. Amsterdam: Dutch University Press.
- Hildebrand, M., Hesper, B.L., Spreen, M., e.a. (2005). *De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie. Een onderzoek naar de betrouwbaarheid en predictieve validiteit van de HCR-20, HKT-30 en PCL-R*. Utrecht: Expertisecentrum Forensische Psychiatrie.
- Hildebrand, M., Spreen, M., Schönberger, H.J.M., e.a. (2006). *Onttrekkingen tijdens verlof, ontvluchtingen en recidives tijdens de tbs-behandeling in de jaren 2000-2005*. Utrecht: Expertisecentrum Forensische Psychiatrie.
- Hudson, S.M., Wales, D.S., Bakker, L., e.a. (2002). Dynamic risk factors: The Kia Marama Evaluation. *Sexual Abuse: A Journal of Research and Treatment*, 14, 103-119.
- Kraan, J. van der, Schekman, M., Nijman, H., e.a. (2008). *Behandelbaarheid bij TBS-patiënten*. *Sancties*, 4.
- Lammers, S.M.M. (2007). Blijven stoute jongens stout? Taxatie van het recidiverisico in de tbs. *De Psycholoog*, 4, 194-201.
- McGuire, J., & Hatcher, R. (2001). *Offense-focused problem solving: preliminary evaluation of a cognitive skills program*. *Criminal Justice and Behavior*, 28, 564-587.
- Nieuwenhuizen, Ch. van, & Philipse, M.W.G. (2002). *Risicotaxatie bij zedendelinquenten: een globaal literatuuroverzicht*. *Tijdschrift voor seksuologie*, 26, 70-78.
- Nijman, H.L.I., de Kruyk, C., & van Nieuwenhuizen, Ch. (2004). *Behavioral changes during forensic psychiatric (TBS) treatment in the Netherlands*. *International journal of law and psychiatry*, 27, 79-85.
- Philipse, M., Koeter, M., van der Staak, C., e.a. (2006). *Static and dynamic patient characteristics as predictors of criminal recidivism: a prospective study in a Dutch psychiatric sample*. *Law and human behavior*, 30, 309-327.
- Vogel, V. de, de Ruiter, C., van Beek, D., e.a. (2003). *De waarde van gestructureerde risicotaxatie. Een retrospectief empirisch onderzoek bij behandelde seksuele delinquenten*. *Maandblad Geestelijke volksgezondheid*, 58, 9-29.
- Vogel, V. de (2005). *Structured risk assessment of (sexual) violence in forensic clinical practice. The HCR-20 and SVR-20 in Dutch forensic psychiatric patients*. Amsterdam: Dutch University Press.
- Werkgroep Risicotaxatie Forensische Psychiatrie (2002). *Handleiding HKT-30, versie 2002. Risicotaxatie in de forensische psychiatrie*. Den Haag: Dienst Justitiële Inrichtingen.

AUTEURS

E. DE JONGE was als onderzoeker en is thans als gz-psychooloog in opleiding verbonden aan FPC de Oostvaarderskliniek te Almere.

H.L.I. NIJMAN was ten tijde van het onderzoek werkzaam bij FPC de Kijvelanden te Poortugaal. Thans is hij als hoofdonderzoeker verbonden aan de divisie Ortho- en Forensische Psychiatrie van Altrecht en als bijzonder hoogleraar Forensische Psychologie aan de Radboud Universiteit te Nijmegen.

S.M.M. LAMMERS is psycholoog en hoofd van de afdeling Onderzoek, Zorgontwikkeling en Opleidingen van FPC Oldenkotte te Rekken.

Correspondentieadres: drs. E. de Jonge, FPC de Oostvaarderskliniek, Postbus 30.178, 1303 AD Almere.

E-mail: e.dejonge@dji.minjus.nl.

Geen strijdige belangen meegeedeeld.

Het artikel werd voor publicatie geaccepteerd op 29-10-2008.

SUMMARY

Behavioural changes during forensic psychiatric treatment: a multicenter study – E. de Jonge, H.L.I. Nijman, S.M.M. Lammers –

BACKGROUND In forensic psychiatry it is essential that patients' risks of reoffending are assessed as reliably as possible. The risk assessment instrument HKT-30 assesses both static and dynamic risk factors. It is assumed that dynamic risk factors should change when a patient receives treatment.

AIM To find out whether dynamic risk factors changed during the course of forensic psychiatric treatment.

METHOD The HKT-30 was administered 984 times for forensic psychiatric patients in 3 different forensic psychiatric centres, in the period June 2003–November 2006. For 513 patients the instrument was administered at least once, for 313 this was done at least twice and for 158 patients for three years in succession. By subdividing the research group into 'new', 'old' and 'transferred' patients we were able to examine in which phase of treatment the largest changes in HKT-30 scores occurred.

RESULTS More than half the scores for dynamic risk factors, as well as the total score, declined significantly as treatment progressed, but the differences were small in absolute terms. The three subgroups of patients hardly differed from each other with regard to the degree of change.

CONCLUSION The scores for the risk factors, assumed in theory to be changeable, seemed to become lower as the treatment progressed. However, it is not yet certain whether these lower scores were in fact directly linked to a reduction on the risk of reoffending.

[TIJDSCHRIFT VOOR PSYCHIATRIE 51(2009)4, 205-215]

KEY WORDS dynamic risk factors, forensic psychiatry, HKT-30, risk assessment