

WaagSchaal jeugd

De psychometrische kwaliteit van een gestructureerd klinisch risicotaxatie-instrument voor de ambulante forensische psychiatrie

JOAN VAN HORN*
JULIA WILPERT**
MARIEKE BOS***
MARA EISENBERG****
JULES MULDER*****

ABSTRACT

Risicotaxatie is het inschatten van de kans dat delinquenten opnieuw een delict plegen. Een aantal risicotaxatie-instrumenten, waaronder de HCR-20 en SVR-20, heeft zijn intrede in Nederland gedaan, maar bleken ongeschikt voor de ambulante forensische poliklinieken. De Waag, centrum voor ambulante forensische psychiatrie, combineerde daarom een aantal bestaande instrumenten aangevuld met in de praktijk gebleken relevante factoren in twee nieuwe gestructureerde klinische risicotaxatie-instrumenten: WaagSchaal volwassenen en WaagSchaal jeugd. In dit onderzoek werden in een groep van 147 jeugdige delinquenten de interbeoordelaars-betrouwbaarheid en de predictieve validiteit van de WaagSchaal jeugd onderzocht. De resultaten lieten zien dat WaagSchaal jeugd betrouwbaar kan worden gescoord en een goede voorspelling oplevert voor geweldsrecidive twaalf maanden na ambulante behandeling.

Kernwoorden: Gestructureerde klinische risicotaxatie – ambulante forensisch psychiatrie – predictieve validiteit – psychometrische evaluatie

INLEIDING

De aard en omvang van de huidige jeugdcriminaliteit in Nederland vormen een serieuze bedreiging van de maatschappelijke orde, zo wordt gesteld in het beleidsprogramma 'Jeugd Terecht' van het Ministerie van Justitie (Openbaar Ministerie, 2006). Niet zozeer vanwege de toename van het aantal aanhoudingen van jeugdige criminelen (deze zou volgens het CBS rond de 47.000 liggen), maar eerder door de veranderende aard ervan. Er worden bijvoorbeeld minder vermogensdelicten geregistreerd, terwijl het aantal geweldsdelicten stijgt (Jeugd terecht, 2002). Het totale aantal minderjarige verdachten van geweldsdelicten zou in de periode van 1995 tot 2002 met 50% zijn gestegen. Deze trend is ook in andere West-Europese landen waar te nemen. Opmerkelijk is dat het aantal voor mishandeling aangehouden minderjarige meisjes naar verhouding sterk toeneemt. Ook onder jeugdige allochtonen in Nederland steeg het aantal geregistreerde criminele activiteiten: van minder dan 11% in 1998 tot 13,5% in 2002 (Jeugd terecht, 2002). Deze toename was vooral te zien bij Oost-Europeanen, Antillianen, Surinamers, Turken, Marokkanen en andere Afrikanen. Autochtone jongeren zouden vergeleken met allochtone jongeren relatief vaker door de politie als verdachte worden geregistreerd (Blom & Van der Laan, 2006). Afkomst zou echter minder samenhangen met

* Hoofd onderzoek De Waag

** BSc Psychologie/Medewerker onderzoek De Waag

*** Psycholoog, Tot oktober 2007 verbonden aan de Waag als medewerker onderzoek

**** Psycholoog

***** Algemeen directeur De Waag

delictgedrag dan factoren zoals niet-thuiswonend, eenoudergezin en een slechte economische status (Jennissen & Blom, 2006).

Deze ontwikkelingen hebben in toenemende mate geleid tot het besef dat er niet één enkele oorzaak is aan te wijzen, maar dat tal van verschillende risicofactoren kunnen leiden tot uiteenlopende vormen van criminaliteit. Risicofactoren kunnen zich voordoen op verschillende gebieden, zoals in het gezin, op school en in de omgeving van de jongere. Maar ook de leeftijd waarop de jongere zijn eerste delict pleegt, blijkt één van de meest belangrijke voorspellers te zijn voor toekomstig delictgedrag. Onderzoek wees uit dat minderjarigen op steeds jongere leeftijd met politie en/of justitie in aanraking komen (zie bijv. Cottle, Lee & Heilbrun, 2001; Lodewijks, De Ruiter & Doreleijers, 2002). De meeste jongeren die al op zeer jeugdige leeftijd crimineel gedrag vertonen, plegen later niet alleen meer, maar ook steeds ernstiger delicten. Het is daarom van groot belang om vroeg in de criminele carrière van jongeren te kunnen onderscheiden wie een relatief hoog risico heeft om in herhaling te vallen.

1. GESTRUCTUREERDE KLINISCHE RISICOTAXATIE-INSTRUMENTEN

Een manier om een inschatting te kunnen maken van het recidiverisico is door gebruik te maken van gestructureerde klinische risicotaxatie-instrumenten. Dit type instrument heeft de voorkeur boven actuariële risicotaxatie-instrumenten omdat het naast statische (door interventie onveranderbare) factoren ook dynamische risicofactoren bevat (door interventie te beïnvloeden). Ondanks het bewezen belang van risicotaxaties voor de effectiviteit van behandeling (Andrews, 1989; Bonta, 1997) worden risicotaxaties bij jeugdige delinquenten in ambulante forensisch psychiatrische zorginstellingen slechts incidenteel uitgevoerd. Eén van de redenen hiervoor is het ontbreken van gevalideerde gestructureerde klinische risicotaxatie-instrumenten voor deze doelgroep. In Nederland worden weliswaar enkele delictspecifieke risicotaxatie-instrumenten voor jeugdige delinquenten gebruikt, maar die zijn voornamelijk gevalideerd in een residentiële forensische populatie: de *Structured Assessment of Violent Risk in Youth* (SAVRY, Borum, Bartel & Forth, 2002) voor geweldplegers en de *Juvenile Sex Offender Assessment Protocol II* (J-SOAP II, Prentky & Righthand, 2003) voor zedendelinquenten.

In Amerikaans onderzoek werd de predictieve validiteit van de SAVRY (= het vermogen om toekomstig delictgedrag te voorspellen) matig bevonden ($AUC = .70$, McEachran, 2002)¹. De correlatie met vergelijkbare andere instrumenten zou redelijk zijn (Borum e.a., 2002). Recente gegevens over de psychometrische eigenschappen van de Nederlandse versie van de SAVRY (Lodewijks, Doreleijers, de Ruiter & de Wit-Grouls, 2003) bij jeugdige delinquenten in een justitiële inrichting wezen uit dat de voorspellende waarde van de SAVRY voor gewelddadige incidenten in de inrichting, goed is ($AUC = .80$, Lodewijks, Doreleijers, de Ruiter & Borum, in druk). Ook de overeenstemming tussen beoordelaars was goed ($ICC = .74$).

1 AUC staat voor Area Under the Curve en geeft de mate aan waarin een instrument het gemeten concept (in dit geval het recidiverisico) voorspelt. Doorgaans worden de AUC-waarden volgens de volgende richtlijnen geïnterpreteerd: $AUC < .70$ = slechte predictieve validiteit; $.70 \leq AUC \leq .75$ = matige predictieve validiteit en $AUC > .75$ = goede predictieve validiteit.

2. ONTWIKKELING VAN EEN RISICOTAXATIE-INSTRUMENT VOOR DE AMBULANTE FORENSISCHE PSYCHIATRIE

De waarde van de SAVRY heeft zich in de Justitiële Jeugd Inrichtingen (JJI) bewezen. Maar zoals eerder aangegeven zijn er nog geen gevalideerde risicotaxatie-instrumenten beschikbaar voor de poliklinische setting, terwijl er wel behoefte aan is. De Waag, centrum voor ambulante forensische psychiatrie heeft daarom een start gemaakt met het ontwikkelen van een gestructureerd klinisch risicotaxatie instrument voor jeugdige cliënten tot 18 jaar, de WaagSchaal jeugd.

De keuze voor de vorm en inhoud van de risicofactoren die in de WaagSchaal zijn opgenomen, kwam voort uit de resultaten van een risicotaxatie pilotonderzoek dat in 2003 werd verricht bij jeugdige geweld- en zedenplegers bij De Waag (Van Horn, Scholing & Mulder, 2005). De algemene conclusie uit dit pilotonderzoek was dat een combinatie van instrumenten veel beter aansloot op de ambulante behandelpraktijk dan de afzonderlijke instrumenten. De instrumenten die qua inhoud en structuur voor de poliklinische populatie het meest relevant werden geacht, werden gecombineerd in de WaagSchaal jeugd. In tabel 1 staan de instrumenten weergegeven die (deels) in de jeugdversie van de WaagSchaal werden opgenomen.

TABEL 1 SAMENSTELLING DE WAAGSCHAAL JEUGD

WaagSchaal jeugd
- Youth Level of Service\Case Management Inventory (YLS\CMI) ^a
- Structured Assessment of Violence Risk in Youth (SAVRY) ^b
- Juvenile Sex Offender Assessment Protocol Dutch (J-SOAP D) ^c
- Static-99 ^d
- Structured Assessment of Protective Factors for Violence Risk (SAPROF) ^e

- a Hoge & Andrews, 1994. Nederlandse vertaling in voorbereiding: De Ruiter, Van Domburgh, Das & Hildebrand (in voorbereiding).
- b Borum, Bartel en Forth (2002). Nederlandse vertaling: Lodewijks, Doreleijers, de Ruiter en De Wit-Grouls (2001).
- c J-SOAP II. Prentky & Righthand, 2003. Nederlandse versie (J-SOAP D): Bullens, Van Horn, Van Eck en Das (2005).
- d Hanson en Thornton (1999). Nederlandse vertaling: Van Beek, De Doncker en De Ruiter (2001).
- e De Vogel, De Ruiter, Bouman & De Vries Robbé (2007).

In de WaagSchaal jeugd wordt tevens ingespeeld op een aantal kanttekeningen die wij plaatsten bij de gangbare gestructureerde klinische risicotaxatie-instrumenten en de manier waarop risicotaxaties worden verricht met deze instrumenten.

1. Dynamische risicofactoren zijn het meest geschikt als doelen voor interventie, omdat alleen in deze factoren verandering mogelijk is. Eén van de beperkingen van gestructureerde klinische risicotaxatie-instrumenten is echter dat ze - ondanks dat ze dynamisch worden verondersteld – voor een groot deel uit statische items bestaan. Van de 24 risicofactoren van de SAVRY bijvoorbeeld zijn tien (43%) statisch, dus onveranderbaar door interventie.
2. Een andere beperking is dat veranderingen in het risicogedrag bij de meeste instrumenten slechts kan worden weergegeven op een 3-puntsschaal. Door deze beperkte scoringsrange is het maar de vraag of bij herhaalde metingen (na bijvoorbeeld 1 jaar) potentiële (subtiële) veranderingen zichtbaar worden.

3. Onderzoek wees uit dat jongeren vaak meerdere typen delicten plegen. Bekend is bijvoorbeeld dat het plegen van zedendelicten vaak onderdeel uitmaakt van een bredere criminele carrière (Van Wijk & Blokland, 1999; Bullens, Van Wijk & Mali, 2006) en dat 50% van de zedendelinquenten recidiveert met andere dan seksuele delicten (Vandiver, 2006). In recidive-onderzoek wordt daarom aangeraden om niet alleen specifieke recidive (recidive met hetzelfde type delict als het indexdelict) in te schatten, maar ook recidive met andersoortige delicten. Niettemin wordt door- gaans het indexdelict gebruikt om een delinquent te typeren als een gewelds- of zedendelinquent, ongeacht diens criminele voorgeschiedenis en op basis daarvan de risicotaxatie te verrichten met delictspecifieke risicotaxatie-instrumenten, zoals de SAVRY.
4. Een jongere heeft - volgens de richtlijnen van risicotaxatie handleidingen - een ‘matig’ of ‘hoog’ recidiverisico als er sprake is van ‘enig’ of ‘ernstig’ disfunctioneren op een bepaald gebied. In de SAVRY bijvoorbeeld heeft een jongere een matig reci- dive risico als hij in enige mate contact heeft met delinquente leeftijdgenoten en een hoog recidive risico als hij vaak met criminele leeftijdgenoten omgaat. De vraag is hoe risicoverhogend dit is als betrokkene een solopleger is en er geen aanwijzin- gen zijn dat betrokkene’s delictgedrag mede beïnvloed wordt door zijn contacten met criminele jongeren. Ofschoon in wetenschappelijk onderzoek werd aange- toond dat omgang met delinquente leeftijdgenoten significant samenhangt met recidive, hoeft dat op het individuele niveau nog niet zo te zijn, ook al gaat de getaxeerde om met criminelen. De voorspelling van recidive wordt mogelijk nauw- keuriger als voor elk geconstateerd probleemgebied moet worden aangegeven of het samenhangt met delictgedrag (in het verleden en de toekomst).

3. RISICOTAXATIE PROCEDURE VAN DE WAAGSCHAAL JEUGD

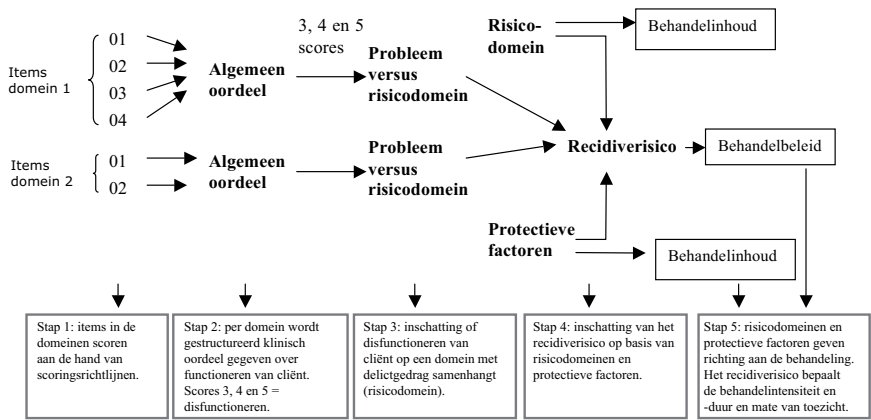
De WaagSchaal jeugd is opgebouwd uit 12 domeinen, elk bestaande uit een aantal items. In tabel 2 staan deze domeinen weergegeven.

TABEL 2 DOMEINEN IN WAAGSCHAAL JEUGD

1.	Eerdere en huidige delicten (8 items)
2.	School/(bij)baan (9 items)
3.	Financiën (3 items)
4.	Woonomgeving (3 items)
5.	Gezin (12 items)
6.	Sociaal netwerk (5 items)
7.	Vrije tijd (3 items)
8.	Middelen (7 items)
9.	Emotioneel/Persoonlijk (16 items)
10.	Houding (5 items)
11.	Risicomanagement (8 items)
12.	Seksuele problematiek (15 items)

De risicotaxatie wordt verricht aan de hand van de volgende vijf stappen (zie Figuur 1):

FIGUUR 1 SCHEMATISCHE WEERGAVE VAN DE STAPPEN IN HET RISICOTAXATIEPROCES WAAGSCHAAL



In stap 1 worden voor elk domein de items gescoord volgens de richtlijnen die in de handleiding van de WaagSchaal staan beschreven. Zo krijgt men een beeld van het functioneren van de cliënt op diverse levensgebieden. Na het scoren van de items wordt op basis van een 6-puntsschaal bij elk domein een gestructureerd algemeen klinisch oordeel gegeven (stap 2). Dit oordeel geeft het algehele functioneren van cliënt weer op het betreffende domein. Leidraad hiervoor zijn de scores op de items en eventuele cliëntspecifieke informatie die niet in de items aan bod is gekomen, maar door de behandelaar wordt meegewogen. De scores 3, 4 en 5 van het algemeen gestructureerd klinisch oordeel geven aan dat de cliënt op het desbetreffende domein in meer of mindere mate onbevredigend functioneert (probleemdomein). Van deze probleem domeinen wordt in stap 3 een inschatting gemaakt van de samenhang met delictgedrag (risicodomeinen). De belangrijkste vraag hierbij is of het disfunctioneren van cliënt op een domein, bijvoorbeeld opleiding/werk, delictgedrag in de hand werkt. Het identificeren van individu-specifieke risicodomeinen is een essentiële stap in de risicotaxatieprocedure omdat deze domeinen richting geven aan de inhoud van de behandeling. Deze werkwijze wordt ondersteund door het interventiemodel van Andrews, Bonta en Hoge (1990). Eén van de drie basisprincipes uit dit model is het 'zorgbehoefteprincipe'. Dit principe gaat er van uit dat de recidivekans bij elk individu (mede) wordt beïnvloed door een specifieke set van criminogene (met delictgedrag samenhangende) factoren. Behandeling, aldus voornoemde auteurs, is het meest succesvol als zij is gericht op deze specifieke set van individuele criminogene factoren.

Tegenwoordig wordt in het verrichten van risicotaxaties ook aandacht besteed aan factoren die een remmende werking hebben op delictgedrag (protectieve factoren). In de WaagSchaal zijn ook protectieve factoren opgenomen omdat daarmee de risicotaxatie veel meer in balans is. Deze protectieve factoren zijn overgenomen uit de *Structured Assessment of Protective Factors for Violence Risk* (SAPROF; De Vogel, De Ruiter, Bouman & De Vries Robbé, 2007). Voor het identificeren van protectieve factoren wordt dezelfde werkwijze gevolgd als bij de risicofactoren: er wordt eerst op itemniveau gekeken op welke gebieden de cliënt goed functioneert en daarna wordt ingeschat of dit voor de individuele cliënt beschermend werkt (key) of kan werken (doel). De risicodomeinen en de protectieve factoren vormen vervolgens samen de input voor het inschatten van het recidiverisico (stap 4). Met deze vier stappen is de fundering gelegd voor het opstellen

van het behandelplan en het nemen van beslissingen over de duur en intensiteit van de behandeling (stap 5).

Ten behoeve van het valideringsonderzoek naar de WaagSchaal jeugd werden de volgende drie vragen geformuleerd:

- 1, Wat is de interbeoordelaarsbetrouwbaarheid voor de verschillende stappen van het risicotaxatieproces van de WaagSchaal jeugd?
- 2, Hoe goed zijn de verschillende fasen van het risicotaxatieproces van de WaagSchaal jeugd in het voorspellen van recidive 12 maanden na behandeling (predictieve validiteit)?
- 3, Voorspelt de WaagSchaal jeugd recidive 12 maanden na behandeling beter dan de SAVRY?

4. METHODE

Het onderzoek was retrospectief van aard en werd uitgevoerd bij De Waag, centrum voor ambulante forensische psychiatrie en onderdeel van de Van der Hoevenstichting. De Waag biedt behandelingen voor mensen met (dreigend) grensoverschrijdend gedrag of strafbaar gedrag. Er wordt met verschillende behandeltypen gewerkt, zoals individuele gesprekstherapieën, partnerrelatietherapieën en groepstherapieën. Groepsbehandelingen zijn er voor plegers van agressiedelicten, huiselijk geweldsdelicten en zedendelicten. Behandeling is cognitief-gedragstherapeutisch van aard en heeft als centrale doel recidive te voorkomen, dan wel te verminderen. De Waag kent een aantal algemene contra-indicaties voor behandeling, waaronder (ernstige) verslavingsproblematiek en het onvoldoende beheersen van de Nederlandse taal. In deze gevallen worden cliënten verwezen naar desbetreffende gespecialiseerde hulpverlening. Eventueel wordt kort daarna gestart met behandeling bij De Waag, al dan niet parallel aan de behandeling van bijvoorbeeld verslaving.

Onderzoeksgroep

De onderzoeksgroep bestond uit 147 jeugdige delinquenten (137 jongens en 10 meisjes) die tussen 2001 en 2002 waren geaccepteerd voor behandeling bij De Waag in verband met het plegen van een (seksueel) geweldsdelict. De gemiddelde leeftijd van de jongeren bij aanvang van de behandeling was 15 jaar ($sd. = 1.55$. Range: 10-18 jaar). De meeste jongeren werden in Nederland geboren (85%). Het geboorteland van de rest van de jongeren was: Nederlandse Antillen (2.8%), Suriname (2.1%), Marokko (2.1%), Irak (2.1%) en overige landen (5.9%), waaronder Somalië en Zuid-Afrika. Het merendeel van de jongeren had slechts de basisschool afgerond (73%) en 10% had op het speciaal onderwijs gezeten. De rest (17%) had naast de basisschool ook lager of middelbaar beroepsopleiding gevolgd.

Behandeling werd bij 100 (68%) jongeren opgelegd door de rechter, 11 (7.5%) waren in afwachting van zitting en 35 (23.8%) namen zonder juridische verplichting deel aan de behandeling. Ruim de helft (59%) maakte de behandeling af. Van de totale groep bleken 72 jongeren eerder met de politie in aanraking te zijn geweest in verband met het plegen van één of meerdere delicten: 75% geweldsdelict (fysiek of verbaal geweld of vandalisme), 38% met een seksueel delict, 39% met een vermogensdelict en 3% met andere delicten, waaronder wapenbezit.

Procedure

De risicotaxaties werden gescoord op basis van dossierinformatie en verricht door getrainde beoordelaars. Om de overeenstemming in scores te berekenen tussen beoordelaars werd de risicotaxatie bij 50 van de 147 jongeren verricht door twee getrainde beoordelaars.

WaagSchaal jeugd

De WaagSchaal jeugd (Van Horn e.a., 2008) is een gestructureerd klinisch risicotaxatie-instrument waarmee op systematische wijze het recidiverisico kan worden ingeschat bij jeugdige delinquenten tot 18 jaar die zijn geïndiceerd voor ambulante forensisch psychiatrische behandeling. Het instrument bestaat uit 94 items verdeeld over 12 domeinen: '1. Eerdere en huidige delicten' (8 items), '2. School/(bij)baan' (9 items), '3. Financiën' (3 items), '4. Woonomgeving' (3 items), '5. Gezin' (12 items), '6. Sociaal netwerk' (5 items), '7. Vrije tijd' (3 items), '8. Middelen' (7 items), '9. Emotioneel/Persoonlijk' (16 items), '10. Houding' (5 items), '11. Risicomanagement' (8 items) en '12. Seksuele problematiek' (15 items). De items worden dichotoom (ja/nee) gescoord of op een 3-puntsschaal (0=niet aanwezig, 1=enigszins aanwezig, 2= aanwezig). De algemeen gestructureerde klinische oordelen aan het eind van elk domein worden gescoord op een 6-puntsschaal van 0 'zeer bevredigend' tot 5 'zeer onbevredigend'. In de WaagSchaal jeugd is de SAPROF opgenomen om de beschermende factoren in kaart te brengen. De items in de SAPROF worden gescoord op een 3-puntsschaal (0=niet aanwezig, 1=enigszins aanwezig, 2= aanwezig). Het recidiverisico wordt ingeschat voor alle delicttypen die in de voorgeschiedenis zijn genoteerd. Het recidiverisico wordt geclassificeerd in termen van 'laag', 'laag-matig', 'matig', 'matig-hoog' of 'hoog'.

Recidive

Recidive werd gedefinieerd als elke nieuwe aangifte van een geweldsdelict (verbaal of fysiek geweld of vandalisme). Informatie over nieuwe aangiften werd verkregen uit het Herkenningsdienstsysteem (HKS) van de politie. Het Herkenningsdienstsysteem (HKS) is een landelijk dekkend systeem dat sinds 1996 door de politie wordt gebruikt om aangiften van misdrijven tegen geïdentificeerde verdachten van 12 jaar en ouder te registreren. Per pleegdatum kunnen meerdere wetsartikelen vermeld staan omdat er op één pleegdatum meerdere delicten kunnen zijn gepleegd, er meerdere wetsartikelen binnen één delict kunnen zijn overtreden en/of de politie de verdenking zo breed mogelijk wil houden. Uit het HKS werden maximaal tien delicten in de voorgeschiedenis genoteerd voorafgaand aan het indexdelict en van de recidivedelicten werden maximaal tien delicten genoteerd vanaf de datum waarop de behandeling werd beëindigd. Van alle jongeren werden de geweldsdelicten tot 12 maanden na behandeling meegenomen.

Analyses

De mate van overeenstemming tussen beoordelaars werd berekend met de Intraklasse Correlatiecoëfficiënten (ICC). Uitgaande van de richtlijnen van Shrout en Fleiss (1979) werden de *single measure* ICC's berekend voor het 'two-way random effect model met een 'absolute overeenstemming' als criterium. De *single measure* ICC-waarden worden

weergegeven om zicht te krijgen op de betrouwbaarheid van de taxatie als deze door een enkele beoordelaar zou worden uitgevoerd. Bij de interpretatie van de ICC's worden doorgaans de volgende richtlijnen aangehouden: $ICC < .40$ = slecht; $.40 \leq ICC < .60$ = matig; $.60 \leq ICC < .75$ = goed; $ICC \geq .75$ = uitstekend (Landis & Koch, 1977).

De recidivepercentages werden berekend met Kaplan-Meier survivalanalyse. Dit is een statistische methode om voor een groep of subgroepen de recidivepercentages en het verloop in tijd te berekenen van het percentage dat recidiveert. In deze berekening wordt gecorrigeerd voor individuele verschillen in de tijdsintervallen tussen het beëindigen van de behandeling en de peildatum. Ieder eerste recidivedelict impliceert een stijging van de recidivecurve; latere recidivedelicten worden niet meegerekend (voor een uitgebreidere uitleg over survivalanalyse zie Wartna, 2000). De Receiver Operating Characteristics (ROC) methode werd gebruikt om de voorspellende waarde (predictieve validiteit) van de WaagSchaal jeugd te berekenen voor recidive met een geweldsdelict in een periode van 12 maanden na afsluiting van de behandeling. Er werd een periode van 12 maanden aangehouden omdat uitspraken over het recidiverisico met gestructureerde klinische risicotaxatie-instrumenten een beperkte 'houdbaarheidsdatum' hebben. Bij jongeren geldt over het algemeen een periode van 6 maanden. Om ROC-analyses te kunnen verrichten moeten er echter voldoende recidivisten zijn. De voorspellingsperiode werd om die reden verlengd naar 12 maanden. Deze periode bleek onvoldoende voor het berekenen van de predictieve validiteit voor recidive met een zedendelict. Er was maar één jongere die met een seksueel delict recidiveerde. De predictieve validiteit werd wel berekend voor recidive met een vermogensdelict. De ROC-analyse levert een waarde op waaraan de voorspellingskracht van het instrument kan worden afgelezen, de zogenoemde AUC-waarde (*Area Under the Curve*). Deze waarde kan minimaal .50 en maximaal 1.00 zijn. Bij de interpretatie van de AUC-waarden worden doorgaans de volgende richtlijnen gehanteerd: $AUC < .70$ = slecht; $.70 \leq AUC \leq .75$ = redelijk; $AUC > .75$ = goed (Sjostedt & Grann, 2002).

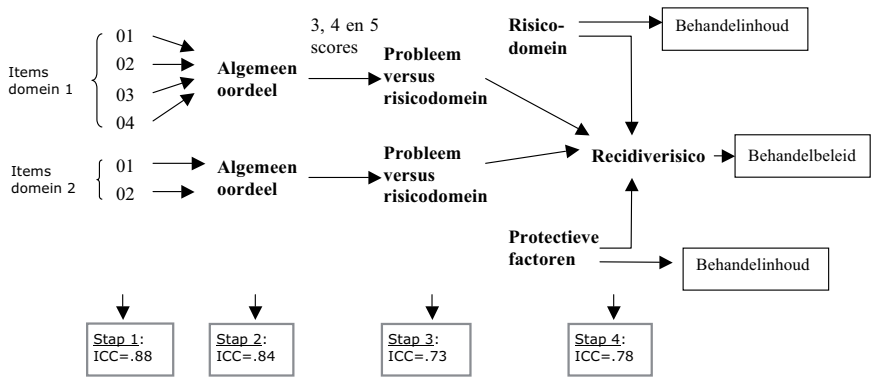
5. RESULTATEN

Onderzoek werd gedaan naar de interbeoordelaarsbetrouwbaarheid en de predictieve waarde van de WaagSchaal jeugd, een gestructureerd klinisch risicotaxatie-instrument voor de ambulante forensische psychiatrie.

Interbeoordelaarsbetrouwbaarheid van de WaagSchaal

Belangrijk bij het verrichten van risicotaxaties is dat de uitkomst ervan zo min mogelijk afhankelijk is van degene die de beoordeling doet. De mate van overeenstemming tussen verschillende beoordelaars werd daarom berekend voor elke stap in de risicotaxatieprocedure van de WaagSchaal (onderzoeksvraag 1).

FIGUUR 2 INTERBEOORDELAARSBETROUWBAARHEID PER STAP IN HET RISICOTAXATIEPROCES WAAGSCHAAL

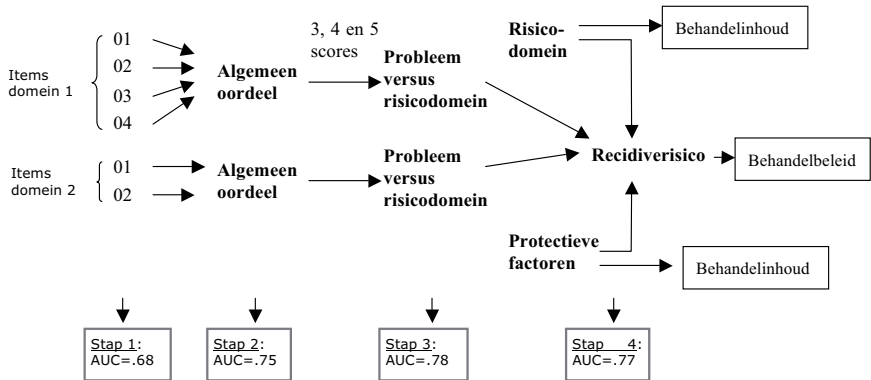


In figuur 2 is te zien dat voor elke stap in het risicotaxatieproces de scores en/of inschatting tussen de twee beoordelaars goed tot uitstekend was. De icc-waarden lagen tussen de .73 voor het identificeren van risicodomeinen (stap 3) en .88 voor het scoren van de items (stap 1).

Recidive en de predictieve validiteit van de WaagSchaal

Kaplan-Meier survival analyses wezen uit dat 12 maanden na behandeling 10.9% van de jongeren recidiveerde met een vermogensdelict, 8.8% met een geweldsdelict en .7% met een seksueel delict.

FIGUUR 3 PREDICTIEVE VALIDITEIT (AUC) VOOR GEWELDSRECIDIVE PER STAP IN HET RISICOTAXATIEPROCES WAAGSCHAAL



Hoe goed voorspellen de diverse stappen in de procedure van de risicotaxatie met de WaagSchaal recidive met een geweldsdelict 12 maanden na behandeling (onderzoeksvraag 2)? Uit figuur 3 blijkt dat de voorspelling van geweldsrecidive op basis van het scoren van items (stap 1) slecht is (AUC=.68). De predictieve waarde bereikte de hoogste AUC-waarde (AUC=.78) in de fase waarin de risicodomeinen worden onderscheiden van de probleemdomeneinen (stap 3).

Voorspellingen met de WaagSchaal jeugd over recidive met een vermogensdelict waren in elke fase slecht (AUC*.60). Aanvullende analyses wezen uit dat van de 16 jon-

geren die met een vermogensdelict recidiveerden, slechts 5 (31%) in de voorgeschiedenis ook een vermogensdelict hadden gepleegd. De rest (69%) had geen criminele antecedenten op dit gebied.

Predictieve validiteit van de SAVRY

Een interessante vraag is natuurlijk hoe de voorspellende waarde van de WaagSchaal jeugd zich verhoudt tot die van andere instrumenten (onderzoeksvraag 3). Als vergelijkingsinstrument werd de SAVRY gescoord en voor de stappen 1 en 4 de predictieve waarden berekend. Stappen 2 en 3 ontbreken in de standaard risicotaxatie-instrumenten. Het scoren van de SAVRY-items (stap 1) leverde een voorspellende waarde op voor geweldsrecidive van $AUC=.67$. Voor de inschatting van het recidiverisico op basis van de SAVRY-items (stap 4) werd een AUC-waarde gevonden van $.68$. Beide waarden vallen in de categorie 'slechte predictieve validiteit'.

6. DISCUSSIE

De WaagSchaal jeugd is een nieuw gestructureerd klinisch risicotaxatie-instrument voor jongeren tot 18 jaar die zijn geïndiceerd voor ambulante forensisch-psychiatrische behandeling. De samenstelling uit (delen van) bestaande risicotaxatie-instrumenten zoals de SAVRY en de J-SOAP D heeft geleid tot een instrument waarmee op een betrouwbare en valide manier recidive kan worden voorspeld. Van de 147 jongeren die voor het plegen van een (seksueel) geweldsdelict waren geïndiceerd voor poliklinische forensisch psychiatrische behandeling recidiveerde 10.9% met een vermogensdelict, 8.8% met een geweldsdelict en .7% met een seksueel delict binnen 12 maanden na afsluiting van de behandeling. De risicotaxatie-procedure van de WaagSchaal kent een aantal opeenvolgend stappen. Eén van de belangrijkste vragen in dit onderzoek richtte zich op de voorspellende kracht van elke stap in deze procedure. ROC-analyses wezen uit dat – voor geweldsrecidive – de predictieve validiteit het beste was in stap 3 waarin de risicodomeinen worden geïdentificeerd. De grootste toename in de voorspelling werd bereikt in stap 2, het gestructureerd klinisch algemeen oordeel. De WaagSchaal bleek minder goed in het voorspellen van recidive met een vermogensdelict. Waarschijnlijk hangt deze uitkomst samen met het feit dat de inschatting van het recidiverisico met een bepaald type delict is gebaseerd op de criminele antecedenten. Als een jongere geen vermogensdelicten in de voorgeschiedenis heeft gepleegd, wordt het risico voor terugval in het plegen van een vermogensdelict niet ingeschat. Van de 16 jongeren die recidiveerden met een vermogensdelict had slechts 31% antecedenten op dit gebied. Alhoewel bij tweederde van de onderzochte jongeren de criminele voorgeschiedenis ontoereikend bleek om recidive met een vermogensdelict te voorspellen, heeft deze informatie een toegevoegde waarde in bijvoorbeeld de typering van een delinquent als zijnde een generalist (iemand die allerlei typen delicten pleegt) of een specialist die zich toelegt op het plegen van één delicttype. Voor de behandeling is deze informatie bijzonder relevant omdat een delinquent die mogelijk vanuit een antisociale persoonlijkheidsstoornis allerlei typen delicten pleegt, meer baat heeft bij een behandeling waarin het kosten-baten aspect meer wordt uitgediept dan aan een behandeling die eenzijdig gericht is op bijvoorbeeld agressieve seksueel-gerelateerde aspecten. Forensisch psychiatrische poliklinieken in Nederland hebben behoefte aan een instrument dat aansluit bij hun populatie. Instrumenten die zijn gevalideerd voor klinische populaties zijn niet zondermeer te gebruiken in een ambulante setting. Dat

bleek ook uit dit onderzoek: de inschatting van het geweldsrecidiverisico met de WaagSchaal jeugd was aanzienlijk beter (AUC=.78) dan met de SAVRY (AUC=.68). Ter vergelijking, bij jongeren in een justitiële jeugdinrichting (JJI) was de predictieve waarde met SAVRY AUC=.86 voor geweldsincidenten in de inrichting en AUC=.71 voor geweldsrecidive 3 jaar na ontslag (Lodewijks, 2008). De WaagSchaal jeugd bleek niet alleen goed in het voorspellen van geweldsrecidive, ook de overeenstemming tussen beoordelaars was goed tot uitstekend in diverse stappen van de risicotaxatie-procedure. In de praktijk bleek dat behandelaars het lastig vonden om in te schatten of een probleemdomein samenhangt met delictgedrag, dus als een risicodomein kan worden beschouwd (stap 3). Toch kwamen ook in deze stap de scores tussen de beoordelaars goed overeen. Dat is een hoopvol resultaat omdat het maken van de vertaalslag van het disfunctioneren op een bepaald domein naar de samenhang daarvan met delictgedrag leidend is voor de inhoud van het behandelplan. Andrews e.a. (1990) bepleiten namelijk dat behandeling van de criminogene factoren, succesvoller is in het terugdringen van recidive dan behandeling van niet-criminogene factoren (d.w.z. probleemdomijnen die niet aan delictgedrag zijn gerelateerd). In een eerdere publicatie wordt de procedure van de risicotaxatie met de WaagSchaal geïllustreerd aan de hand van een casus (Van Horn, Wilpert, Bos & Mulder, 2008). Met het onderhavige onderzoek is de eerste stap gezet in het valideren van de WaagSchaal jeugd. Uiteraard zal vervolgonderzoek moeten uitwijzen of deze positieve bevindingen gerepliceerd kunnen worden in diverse subgroepen jeugdige delinquenten. Ook met de WaagSchaal volwassenen is in september 2008 een eerste valideringsonderzoek gestart. De resultaten worden in de eerste helft van 2009 verwacht.

LITERATUUR

- ANDREWS, D.A. (1989). Recidivism is predictable and can be influenced: using risk assessments to reduce recidivism. *Forum on Corrections Research*, 1, 1-17.
- ANDREWS, D.A., BONTA, J. & HOGE, R.D. (1990). Specification for effective rehabilitation. *Criminal Justice and Behaviour*, 17(1), 19-52.
- BLOM, M. & LAAN, A.M. VAN DER (2006). *Monitor Jeugd terecht 2006*. Factsheet, nr. 2006-4. Den Haag, WODC, 2006.
- BONTA, J. (1997). Do we need theory for offender risk assessment? *Forum on Corrections Research*, 9, 42-45.
- BORUM, R., BARTEL, P. & FORTH, A.E. (2002). Manual for the Structured Assessment of Violent Risk in Youth (SAVRY), consultation edition, version 1. University of South Florida.
- BULLENS, R.A.R., WIJK, A.Ph. VAN & MALL, B. (2006). Similarities and differences between the criminal careers of Dutch juvenile sex offenders and non-sex offenders. *Journal of Sexual Aggression*, Vol. 12, No 2, pp.155-164.
- COTTLE, C.C., LEE, R.J. & HEILBRUN, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior*, 28, 37-394.
- HORN, J.E. VAN, SCHOLING, A. & MULDER, J. (2005). Risicotaxatie bij jeugdige zeden- en geweldplegers in een ambulante setting. *Diagnostiek-wijzer*, 8 (2), 63-76.
- HORN, J.E., WILPERT, J., BOS, M., & MULDER, J. (2008). WaagSchaal jeugd. Risicotaxatie-instrument voor de Ambulante Forensische Psychiatrie. Handleiding. Interne uitgave: De Waag.
- HORN, J. VAN, WILPERT, J., BOS, M.G.N. & MULDER, J. (2008). De WaagSchaal: een andere visie op risicotaxatie. *Panopticon*, 6, 28-39.

- JENNISSEN, R.P.W. & BLOM, M. (2006). *Allochtone en autochtone verdachten van verschillende delicttypen nader bekeken* [Werktitel]. Den Haag: WODC.
- Jeugd Terecht (2002). Actieprogramma aanpak jeugdcriminaliteit 2003-2006.
- LANDIS, J.R. & KOCH, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- LODEWIKS, H.P.B. (2008). Violence risk assessment in adolescents in the Dutch juvenile justice system. Studies on the reliability and predictive accuracy of the Savry. Proefschrift ter verkrijging van de graad Doctor aan de Vrije Universiteit van Amsterdam. Rozenburg Publishers.
- LODEWIKS, H., RUITER, C. de & DORELEIJERS, Th. (2002). Risicotaxatie en risicohantering van gewelddadig gedrag bij adolescenten. *Directieve Therapie*, 23 (1), 25-42.
- LODEWIKS, H.P.B., DORELEIJERS, Th., RUITER, C. de & BORUM, R. (in druk). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry*.
- LODEWIKS, H.P.B., DORELEIJERS, Th., RUITER, C. & WIT-GROULS, H. de (2001). *Gestructureerde taxatie van geweldsrisico bij jongeren* (geautoriseerde vertaling en onderzoeksversie van de SAVRY). Eefde: Rentray.
- McEACHRAN, A. (2002). *The predictive validity of the PCL-YV and the SAVRY in a population of adolescent offenders*. Unpublished master's thesis. Burnaby, British Columbia: Simon Fraser University.
- Openbaar Ministerie (2006). www.justitie.nl/onderwerpen/jeugd/jeugdcriminaliteit
- PRENTKY, R.A. & RIGHTHAND, S. (2003). *Juvenile Sex Offender Assessment Protocol: Manual*. Center for Sex Offender management. Website: www.csom.org.
- SJÖSTEDT, G. & GRANN, M. (2002). Risk Assessment: What is Being Predicted by Actuarial Prediction Instruments? *International Journal of Forensic Mental Health*, 1(2), 179-183.
- SHROUT, P.E. & FLEISS, J.L. (1979). Intraclass Correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86 (2), 420-428.
- VANDIVER, D.M. (2006). A prospective analysis of juvenile male sex offenders. Characteristics and recidivism rates as adults. *Journal of Interpersonal Violence*, 21 (5), 673- 688.
- VOGEL, V. de, RUITER, R. de & BOUMAN, Y., & DE VRIES ROBBÉ, M. (2007). *Handleiding bij de SAPROF. Structured Assessment of PROtective Factors for violence risk. Versie 1*. Intern rapport. Utrecht: Van der Hoevenstichting.
- WARTNA, B. (2000). Recidive-onderzoek en survival analyse. Over het meten van de duur van de misdrijfvrjje periode. *Tijdschrift voor Criminologie*, 42, 2-20.
- WIJK, A.Ph. van, & A.A.J. BLOKLAND (1999). Dé jeugdige zedendelinquent bestaat niet. Een profiel van verkrachters en misbruikers van kinderen. *Proces*, 5/6, 1999, p. 67-70.