

Van gevaar naar risico: regelen in plaats van waarschuwen

Risk assessment in de forensische psychiatrie

H. J. C. VAN MARLE

SAMENVATTING Het beoordelen van risico's bij verlopen en beëindiging van de behandeling bij forensisch-psychiatrische patiënten is zeer belangrijk wegens de grote consequenties voor de maatschappij en voor de patiënt zelf. Het gebruikelijke klinisch onderzoek, uitgevoerd door een psychiater of een psycholoog, blijft afhankelijk van de setting van het onderzoek en de deskundigheid van de onderzoeker. Het verdient daarom aanbeveling naast dit onderzoek een meer objectieve maatstaf in te voeren in het gehele proces van rapportage pro Justitia, gevangenisstraf en tbs-behandeling. Ervaringen met risicobeoordelingsschalen, zoals opgedaan in Canada, stemmen hoopvol. Deze schalen moeten echter nog gevalideerd worden voor de Nederlandse situatie.

Deze objectiverende methode kan niet het individuele onderzoek naar de mate van toerekeningsvatbaarheid vervangen, omdat bij dat onderzoek alleen de hermeneutische methode recht doet aan de specifieke combinatie van psychische stoornis en delict. Ook tijdens de forensisch-psychiatrische behandeling blijft de individuele diagnostiek essentieel voor het behandelingsplan. De verwachting is dat door het gebruik van gestandaardiseerde beoordelingsschalen de behandeling specifiek en meer doeltreffend zal zijn ten aanzien van het hanteren van de risico's van de patiënt.

De implementatie van een risk-assessment-/risk-managementcyclus in het forensisch-psychiatrische circuit verdient aanbeveling om deze patiënten intensief te volgen. De risico's die zij aanvankelijk voor de maatschappij opleveren, kunnen nog lang na het ontslag blijven bestaan.

[TIJDSCHRIFT VOOR PSYCHIATRIE 42(2001)3, 151-161]

TREFWOORDEN forensisch-psychiatrische rapportage, hermeneutiek, risk assessment, risk management

Nog steeds is in de forensische psychiatrie het voorspellen van recidivegevaar het belangrijkste vraagstuk. Dat is niet zo verwonderlijk, aangezien met het inschatten van toekomstige gevaarlijkheid bij psychiatrische patiënten die in contact zijn gekomen met justitie, een belangrijke maatschappelijke functie wordt vervuld. Vooral blijkt die verantwoordelijkheid wanneer zo'n patiënt al eerder is veroordeeld, en er dus de mogelijkheid is geweest hem psychiatrisch te

onderzoeken. De op basis van psychiatrisch of multidisciplinair gedragskundig onderzoek verkregen rechterlijke machtiging, voorwaardelijke straf of terbeschikkingstelling (tbs) blijkt echter niet voldoende om gevaarlijke mensen te weerhouden van hernieuwd (seksueel) geweld. Bij een ambulante behandeling als voorwaarde bij een voorwaardelijk strafdeel of een tbs is de rechterlijke macht grotendeels afhankelijk van de adviezen van psychiaters en gedragskundigen. Het

adviseren tot proefverlof gaat geheel van de behandelaars zelf uit, de rechter of het ministerie beslist op grond van het advies.

Vanwege de ernstige risico's die terbeschikkinggestelden kunnen opleveren voor de maatschappij, is het niet verwonderlijk dat er vanuit de tbs-klinieken grote belangstelling bestaat voor het verbeteren van de risicobeoordeling bij hun patiënten en dat zij verder zijn gaan kijken dan het tot dan toe gebruikte klinisch onderzoek (Hildebrand & De Ruiter 1999; Verhagen & Philipse 1995; Brand e.a. 1998). Aansluiting is gevonden bij de gelijktijdige ontwikkelingen in Canada, waar al langer systematisch onderzoek wordt gedaan naar het beoordelen van mogelijk gevaarlijke personen (Webster e.a. 1985; Webster e.a. 1997). Deze gang van zaken heeft ertoe geleid dat het systematisch beoordelen van het gevaarsrisico, het *risk assessment*, in Nederland bekend is geworden en ook hier wordt onderzocht op zijn praktische bruikbaarheid. Is nu het klassieke psychiatrische onderzoek overbodig of ouderwets geworden? Hoe verhouden dit onderzoek en het *risk assessment* zich tot elkaar? Wat betekent de introductie van *assessment* voor de (justitiële) gezondheidszorg?

In dit artikel wordt nader ingegaan op deze andere wijze van 'diagnosticeren': het (eventueel herhaald) in kaart brengen van bepaalde risicofactoren bij een bepaalde persoonlijkheid om op die wijze gevaar voor anderen te voorspellen.

'PREDICTION'

De ontwikkelingen in het voorspellen van gevaarlijk gedrag gaan langzaam, alhoewel er vorderingen vallen te constateren (Van Marle 1986, 1992). Het denken over gevaar, gevaarlijkheid en gevaarlijk gedrag is veranderd, evenals de wijze waarop deze concepten verder zijn onderzocht. Het voorspellen van een zekere mate van waarschijnlijkheid van optreden van bepaald ongewenst gedrag lukte niet op basis van het gesystematiseerd onderzoeken van de dossiers, de zogenoemde actuariële methode. In theorie

was voorspellen al helemaal niet op een wetenschappelijke wijze mogelijk. Afgezien van de wisselende criteria voor gevaar, gevaarlijkheid en gevaarlijk gedrag en de wisselende diagnostiek van het (psychiatrische/psychologische) toestandbeeld van de daders blijkt een statistische voorspelling alleen al niet haalbaar door de lage frequentie (de *base rate*) van het te voorspellen gevaarlijke gedrag. Hoewel bijvoorbeeld het voorspellen van geweld binnen het gezin mogelijk is vanwege het frequente voorkomen ervan, is het voorspellen van een moord, laat staan partnerdoding, dat niet (Campbell 1995). De incidentie ervan is gewoon te laag. In deze relatief zeldzame gevallen is het echter wel goed gebruik om de kans op herhaling te beoordelen, omdat de daad op zichzelf genomen zeer ernstig is en een forse inbreuk op de rechtsorde betekent. Dergelijke gevallen kunnen dus alleen op zichzelf worden beoordeeld.

Het statistisch opsporen van een mogelijke dader in een populatie brengt een onaanvaardbaar hoog aantal ten onrechte als dader aangemerkte personen met zich mee, zelfs als de gebruikte test – die overigens niet bestaat – de wetenschappelijk aanvaarde waarschijnlijkheidskans van 5% zou kennen. Niet alleen brengt een lage *base rate* een hoog aantal fout-positieven met zich mee en dus een lage specificiteit ten opzichte van het criterium (Blazer & Hays 1998), maar ook moet een test nog gevalideerd worden voordat hij binnen een bepaalde populatie gebruikt kan worden: is de sensitiviteit (het vaststellen van de ware positieven) gelijk bij een Canadese gevangenispopulatie en bij de Nederlandse tbs-gestelden? Het gebrek aan nauwkeurigheid brengt met zich mee dat, afhankelijk van de gebruikte methode, er altijd een zeker aantal onschuldigen zal worden aangemerkt als dader, en daardoor bijvoorbeeld ten onrechte zal worden opgesloten. Dit risico is dus inherent aan het voorspellen van gevaarlijk gedrag en zal op zijn maatschappelijke consequenties meegewogen moeten worden bij het gebruik van elke methode.

Een ander probleem, dat het meest zichtbaar is en sterk bijdraagt aan de discussie over het inschatten van gevaar, is dat van de fout-negatieve conclusies: iemand wordt beoordeeld als niet gevaarlijk maar blijkt dat wel te zijn. Echter, de positieve en de negatieve voorspellende waarde van een test hangen met elkaar samen (Blazer & Hays 1998). Wanneer een testuitkomst geen fout-negatieve uitspraak meer mag bevatten, wil dat zeggen dat een sensitiviteit van 100% moet worden bereikt door alle scores die wijzen op mogelijke gevaarlijkheid als zonder meer positief voor gevaar aan te merken. In dat geval neemt het aantal fout-positieven weer toe: die zijn weliswaar dragers van een of meer risicofactoren, maar daarnaast bestaat er bij hen ook een groot aantal niet-*risicofactoren* of beschermende (*protective*) factoren. De voorspellende waarde van een test wordt bepaald door de plaats van de afkapwaarde: die verhouding tussen de sensitiviteit en de specificiteit waarbij het voorspellend vermogen optimaal is. Zoveel mogelijk hits (ware positieven) en zo weinig mogelijk misses (fout-negatieven), met een aanvaardbaar aantal *false alarms* (fout-positieven). Over afkapwaarden valt derhalve te onderhandelen.

Niet alleen de methode, maar ook de termijn waarover een voorspelling wordt gedaan, blijft een bron van onzekerheid. Na een uitgebrachte voorspelling kan er immers in een bepaalde periode veel veranderen in en rondom de dader. Hij kan jaren in een gevangenis hebben gezeten, meer of minder intensieve hulpverlening hebben gehad, werk hebben gehad, een relatie hebben opgebouwd – allerlei omgevingsafhankelijke factoren hebben invloed op het al of niet herhalen van crimineel gedrag door een mogelijke dader (Haen Marshall 1998; McGuire 1995). Vergelijkbare beoordelingen doen op verschillende momenten maakt de veranderingen van en rondom de persoon zichtbaar in de loop der tijd, waarbij het er niet toe doet hoe die veranderingen tot stand zijn gekomen. Het is niet bekend of die factoren die blijken te voorspellen bij bekende daders (de recidivisten), ook gelden

bij *first offenders*. Voorspellingen doen over iemand die al een misdrijf heeft gepleegd, is eenvoudiger dan voorspellingen doen over iemand die (nog) geen justitiecontacten heeft gehad. De eerste is immers bekend, inclusief zijn gevaarlijk gedrag dat kan worden geanalyseerd, en eerder gevaarlijk gedrag is een goede voorspeller voor herhaling ervan.

GEDRAGSKUNDIG ONDERZOEK

Het medisch-psychiatrisch onderzoeken van verdachten van een misdrijf is al zo oud als de geneeskunde zelf. Het op het individu gerichte medisch onderzoek geeft immers, afhankelijk van de stand van zaken in de medische wetenschap, een goed beeld van de gevoelens en beweegredenen van de onderzochte met betrekking tot het tenlastegelegde. Sedert enkele decennia vindt ook veel individueel gericht onderzoek plaats door een psycholoog, al dan niet in combinatie met een psychiater. Zo'n psychiatrisch/psychologisch onderzoek heeft een aantal kenmerken. Ten eerste is er de deskundigheid van de onderzoeker, die vanuit zijn vakgebied rapporteert aan de instantie die om het onderzoek heeft verzocht en die ook de vraagstelling heeft opgesteld. Naast de gebruikte psychiatrische en psychologische methodiek speelt de persoonlijke interpretatie van de gegevens door de onderzoeker een rol. Ten tweede is het onderzoek bekend bij de aanvragers: de methode ervan is begrijpelijk en tot op zekere hoogte controleerbaar. Ten derde is het onderzoek maatwerk, want het is toegespitst op de onderzochte zelf en de voor diens persoon specifieke vraagstelling. De aanvrager is meestal tevens degene die met het onderzoek verder zal gaan werken. Door de specifieke vraagstelling zijn de gegevens niet zonder meer te gebruiken voor de beantwoording van andere vragen of voor andere doeleinden. De opbouw en de conclusie worden immers geleid door de vraagstelling. Ten vierde wordt de tijd die nodig is efficiënt gebruikt, aangezien de onderzoeker aanwezig is zolang het onderzoek duurt.

Tevens kan diens inzet afgestemd worden op de ernst van het misdrijf en de vraagstelling van de rechtbank.

Deze eigenschappen hebben echter nadelen. Er is een expert nodig, maar wie beslist over diens deskundigheid in het algemeen en ten aanzien van de vraagstelling in het bijzonder? Elke onderzoeker is vrij om op zijn eigen wijze te rapporteren, waardoor er geen systematische weergave voor alle onderzoeken wordt verkregen. Tevens is er sprake van een persoonsgebonden oordeel, dat door de lage graad van vergelijkbaarheid van het onderzoeksrapport moeilijk te controleren valt. Twee deskundigen kunnen op basis van hun eigen onderzoek het nogal eens geheel of gedeeltelijk niet eens zijn met de conclusies van hun collega, waardoor bij de aanvragers als niet-deskundigen twijfel rijst aan de betrouwbaarheid van beider conclusies. Een derde, onafhankelijk oordeel is dan nodig, maar in principe lijdt dat aan dezelfde tekortkomingen.

Naarmate gedragskundig rapporteren steeds meer voorkomt en er grote belangen op het spel staan, wordt er kritischer naar de rapportages gekeken en al snel tot een contra-expertise of *second opinion* besloten. Deskundigen komen al dan niet persoonlijk rechtstreeks tegenover elkaar te staan en de tekortkomingen in de gebruikte gedragskundige methodes worden pijnlijk zichtbaar. In toenemende mate wordt er nu gevraagd om een meer feitelijk geconstrueerde, empirisch gebaseerde en gestandaardiseerde onderzoeksmethode die, voor het moment zelf en over een langere periode, meer vergelijkingsmogelijkheden biedt voor individueel onderzoek. Een mogelijke oplossing is gevonden in de constructie van risicoschalen in de vorm van vragenlijsten, het zogeheten *risk assessment* (Monahan & Steadman 1994).

RISK ASSESSMENT

Het voorspellen van gevaar op basis van dossiergegevens alleen, de *actuarial prediction*, is praktisch niet bruikbaar gebleken wegens de moeiza-

me definiëring van criteria, het sterk situatief bepaald zijn van gevaarlijk gedrag, en de methodologische bezwaren (Quinsey e.a. 1998). Vandaar dat de inspanningen om gevaar voor anderen te voorkomen, zich niet zozeer meer moeten richten op het voorspellen van gevaarlijk gedrag, maar op de risico's die iemand oplevert op basis van bepaalde persoonlijke en situatieve factoren (Litwack e.a. 1993). De uit empirisch onderzoek gebleken gevaarsvoorspellers, 'predictoren', zijn gebundeld tot biopsychosociale domeinen waarbinnen risicofactoren, voorzover momenteel bekend, kunnen optreden en zijn te registreren. Steadman e.a. (1994) hebben deze factoren op een rij gezet voor hun onderzoek. De vier domeinen die zij gebruiken, zijn: (1) (aangeboren) eigenschappen, in het bijzonder persoonlijkheidsfactoren en cognitieve factoren, en tevens geslacht, milieu van herkomst en familiale belasting; (2) de voorgeschiedenis, in het bijzonder de sociale wordingsgeschiedenis (gezin, opleiding en werk), de psychiatrische en de criminologische voorgeschiedenis; (3) situatieve factoren, in het bijzonder de subjectief ervaren stress, de maatschappelijke ondersteuning, en de gebruikte geweldsmiddelen; (4) psychiatrische kenmerken, zoals de DSM-IV-As-I- en As-II-diagnose, de aanwezigheid en ernst van belastende symptomen (wanen, hallucinaties, gewelddadige fantasieën), middelengebruik en mate van psychosociaal functioneren.

Retrospectief is het mogelijk bij daders allerlei voorspellende factoren vast te stellen, die zich zelfs tot voor de geboorte kunnen uitstrekken (Raine 1993). Of deze factoren prospectief herleid kunnen worden tot een of ander soort gevaarlijk gedrag valt nog te bezien. Vele kunnen in de loop van het leven gecompenseerd worden door opvoedings- en milieufactoren. Van veel mensen zullen we nooit weten of die zulke belastende factoren bij zich dragen: zij worden ons niet bekend. Wanneer bepaalde risicofactoren wel bekend zijn, is het vervolgens nog maar de vraag of die veranderd kunnen worden. Sommige, zoals geslacht, sociaal-economische

klasse, de leeftijd waarop de eerste veroordeling heeft plaatsgevonden en eerdere veroordelingen, staan vast of zijn in het verleden verankerd en kunnen niet worden veranderd. Zij heten daarom statische factoren, in tegenstelling tot dynamische factoren, die nog wel door interventies kunnen worden beïnvloed. De onder 1 en 2 genoemde factoren zijn typisch statisch, zij kunnen niet veranderd worden. De onder 3 en 4 genoemde echter wel, bijvoorbeeld door een behandeling of een begeleidingsprogramma: dit zijn dynamische factoren.

Assessment heeft dus een aantal kenmerken: (1) er wordt gewerkt met vragenlijsten en schalen waarop de antwoorden worden afgezet; (2) deze schalen zijn vergeleken en getoetst op bepaalde onderzoeksgroepen en ontlenen daaraan hun standaard; zij zijn dus onderling vergelijkbaar: 'empirisch toetsbaar'; (3) de methode is objectief: de vragen zijn altijd dezelfde en worden op dezelfde manier gesteld; de beoordeling van de antwoorden gebeurt met de standardscore voor de verschillende onderzochte groepen, zoals geweldsdelinquenten, pedoseksuelen; (4) de informatie is gestructureerd verkregen en is beoordeeld op volledigheid; (5) de informatie is eenduidig qua beantwoording.

Aan deze methodiek kleven echter ook bezwaren. Zo zijn er specialisten nodig, speciaal opgeleid in het uitwerken van de tests, om de scores van de vragenlijsten om te zetten in bepaalde uitspraken over de persoon. De validering van de vragenlijsten bij bepaalde groepen is bepalend voor de afname bij andere groepen, die bepaalde vragen wel eens heel anders kunnen interpreteren ('meet de lijst nog wel wat deze pretendeert te meten?'; de interne en externe validiteit). Validering binnen de gewenste populatie, die ook weer zeer heterogeen kan zijn, zoals bijvoorbeeld de Nederlandse tbs-gestelden, is noodzakelijk. Bij deze validering horen ook de afkappunten: dit wil zeggen dat het arbitrair is wanneer een zekere mate van gevaarlijkheid, deviatie of ziekte wordt aangenomen; de omslagpunten berusten op afspraken tussen deskundigen onderling of tussen

deskundigen en beleidsmakers. Een lijst moet verder een voldoende aantal dynamische factoren bevatten, wil hij geschikt zijn om er de behandelingsvoortgang aan af te meten, anders blijft iemands prognose permanent afhankelijk van zijn (onveranderbare) verleden. Een nadeel is dat het gebruik van deze lijsten tamelijk onbekend is bij behandelaars, directeuren en beleidsmakers, waardoor ze worden ondergewaardeerd.

Er wordt echter wel winst geboekt door de risicofactoren bij één persoon te inventariseren en die te vergelijken met de op dezelfde wijze verzamelde factoren bij de gehele groep. Het risicogedrag van deze persoon wordt voorspeld door de gegevens van diens assessment af te zetten tegen die van de groep waartoe hij behoort. Daar deze groep, de referentie, inmiddels al jaren verder is, is hun geneigdheid tot herhaald risicogedrag bekend op basis van bijvoorbeeld politiegegevens. Risk assessment verschaft op die grond nauwkeuriger en meer betrouwbare resultaten dan het klinisch onderzoek. In het bijzonder worden er minder fouten gemaakt van het fout-positieve en fout-negatieve type (Gardner e.a. 1996) en blijken criminologische variabelen veel betere voorspellers dan klinische (Bonta e.a. 1998). Hierbij mag echter niet vergeten worden dat het nog steeds gaat om risicopercentages, de waarschijnlijkheid van risico's dus. De mogelijkheid bestaat dat een individueel geval niet wordt opgemerkt.

HERMENEUTIEK

Wanneer uit grote empirische onderzoeken, zoals de meta-analyses (Rosenthal 1995) van vele andere gepubliceerde artikelen (Bonta e.a. 1998; Hanson & Bussière 1998), blijkt dat gevalideerde vragenlijsten een meer betrouwbare risicovoorspelling geven dan het gebruikelijke klinisch onderzoek, is het dan nog wel verantwoord om op basis van klinisch onderzoek beslissingen te nemen? Het klinisch onderzoek is immers uiteraard hermeneutisch van aard, dat wil zeggen de informatie van het onderzoeksgesprek interpre-

terend. De onderzoeker met diens persoonlijkheid speelt er een bepalende rol in. De geringe standaardisering en reproduceerbaarheid ervan en de onmiskenbare subjectiviteit vormen de bezwaren waardoor de assessmentmethode zoveel veld heeft kunnen winnen. Aan de hermeneutiek is echter ook veel te danken. Deze methode heeft het ons mogelijk gemaakt de patiënt te begrijpen en zij is onmisbaar voor het ontwikkelen van hypothesen omtrent iemands beweegredenen (Denzin & Lincoln 1994), zoals ten aanzien van de motieven voor een bepaald delict. In bredere zin maakt hermeneutiek het ook mogelijk hypothesen op te stellen op basis van ervaringsfeiten, zoals angst en schuldgevoel, en die vervolgens te toetsen aan een algemene zingeving van menselijk denken, voelen en handelen (Mooij 1998). Bij het vaststellen van de doelen van de verschillende psychotherapieën, zowel de psychodynamische als de cognitieve, is zij zelfs de enige manier waarop de patiënt en de therapeut met elkaar aan het werk kunnen.

Geaccepteerd moet worden dat de hermeneutiek niet meer de enige manier is om tot een goede inschatting en diagnostiek te komen. Diagnostische schalen zoals de *Present State Examination*, het *Structured Clinical Interview for DSM of de Symptom Checklist*, zijn inmiddels ingeburgerd bij het onderzoek van individuen, naast de vele psychologische tests. Voor het forensisch-psychiatrisch onderzoek met betrekking tot de gebrekkige ontwikkeling en ziekelijke stoornis van de geestvermogens is psychologische testdiagnostiek in toenemende mate belangrijk in verband met de pathologische persoonlijkheidsfactoren en de toekomstige gevaarlijkheid. Zowel de onderzochte als de rechtbank heeft recht op *present state of the art* in de diagnostiek (Van Marle 1999a). De beoordeling van de mate van toerekeningsvatbaarheid vereist wegens de unieke combinatie van de verdachte en het hem ten laste gelegde delict enerzijds de kwalitatieve, hermeneutische methode, maar anderzijds – denk aan de rechtsgelijkheid en de zorg op maat (Van Marle 1999b) – zouden de verschillende gradaties

van toerekeningsvatbaarheid wegens hun grote en uiteenlopende consequenties beter en reproduceerbaar van elkaar onderscheiden moeten worden. Bij de beoordeling van het recidivegevaar van een tbs-gestelde is een objectiverende maatstaf naast het klinisch-diagnostisch vervolgonderzoek op verschillende momenten van de tbs zeer gewenst. De rechter heeft voor zijn oordeelsvorming voor het vonnis de behoefte aan rapportage omtrent de verdachte op het moment van het delict en omtrent diens achtergrond, maar vervolgens spelen er grote maatschappelijke belangen met betrekking tot recidive. Ook ten aanzien van die toekomstige risico's moet de forensisch-psychiater zich uitlaten. De rechter kan dit oordeel meewegen in zijn vonnis; de forensisch-psychiatrische behandeling richt zich dan in het bijzonder op het voorkómen van recidive.

DE GESTANDAARDISEERDE ONDERZOEKSMETHODE

Met betrekking tot de inschatting van toekomstige risico's zal zowel het individu als de maatschappij recht gedaan moeten worden. Dat wil zeggen een goede individuele risicobeoordeling en de mogelijkheid om nadien na te gaan of de toegepaste begeleiding of therapie heeft geholpen of niet. Het assessment met behulp van vragenlijsten is daarbij noodzakelijk om in de loop der tijd vergelijkbare uitkomsten te krijgen. Zo kan niet alleen bij het onderzoeksmoment voor de terechtzitting risk assessment worden gedaan, maar ook bijvoorbeeld bij de opname in een tbs-kliniek, bij de zesjaarsevaluatie van de tbs en bij een mogelijke indicatie voor de *long-stay*, de verblijfsafdeling van blijvend gevaarlijke tbs-gestelden. Het hoeft niet zo te zijn dat alle onderzoeksinstrumenten op elk moment worden afgenomen, maar er moet wel een vaste kern van relevante gegevens van meet af aan aanwezig zijn, bijvoorbeeld ten aanzien van vijandigheid, impulscontrole of seksuele deviatie.

Bij de beoordeling van de testuitslagen is de

hermeneutiek weer van belang om samen met de klinische diagnostiek tot een behandelplan en de bijstelling daarvan in de loop der tijd te komen. Een behandelplan moet namelijk op de patiënt zijn toegesneden en voor hem duidelijke motiverende factoren bevatten. Ook zullen niet alle mogelijke therapieën tegelijkertijd uitgevoerd kunnen worden en zal er een zekere hiërarchie in moeten worden aangebracht. Inzicht in de betekenisgeving van de patiënt aan zijn leefwereld en diens mogelijkheden en beperkingen op grond van de psychiatrische diagnostiek kunnen hierbij niet worden gemist. Vervanging van het klinisch diagnostisch onderzoek door uitsluitend assessment-schalen geeft daardoor onnauwkeurigheden bij de indicatiestelling ('niet alles is voor iedereen geschikt') en bij de behandelplanning ('wat op het ene moment nog niet haalbaar is, is dat op een ander moment wel').

Een combinatie van de hermeneutische methode en de beoordelingsschalen (assessment) geeft mijns inziens de beste resultaten, omdat enerzijds de patiënt goed in beeld wordt gebracht, en anderzijds een herhaalbare, objectiverende toets wordt aangebracht. Deze combinatie zal ik verder de gestandaardiseerde onderzoeksmethode noemen. De voordelen ervan zijn aldus de positieve kanten van de beide eerder genoemde methoden: onderling vergelijkbare beoordelingen; duidelijke en consistente risicobeoordeling; beter geïnformeerd en vollediger klinisch oordeel; onderzoeks- en besluitvormingsproces worden zichtbaar gemaakt; specifieke zorg en specifieke patiënten worden gedefinieerd; methode is herkenbaar voor stafleden en zorgmanagers; onnodig insluiten en ten onrechte vrijlaten worden voorkomen; kritiek op de organisatie wordt voorkomen; wetenschappelijke evaluatie wordt mogelijk gemaakt; concrete afspraken over doelstellingen zijn te maken tussen behandelaren en managers.

RISK MANAGEMENT

De onderzochte wordt met de gestandaardiseerde

onderzoeksmethode niet alleen beoordeeld naar de maatstaven van de onderzoekende gedragskundigen, maar ook naar die maatstaven zoals die uit empirisch onderzoek over misdaad, maatschappij en geestelijke stoornis naar voren zijn gekomen. Hoe meer risicofactoren aanwezig zijn, hoe groter de kans is dat iemand een (herhalings)delict zal plegen – een constatering die vooral van belang is voor de praktijk: personen met een laag risico zijn op die manier te identificeren. Het selecteren van personen met een hoog risico blijft veel moeilijker vanwege mogelijke beschermende factoren (Grubin 1999).

Het risk assessment van iemand zal vervolgens moeten leiden tot *risk management*: de risicogebieden, zoals aangegeven op de verschillende lijsten, leveren gegevens op die leiden tot een concrete aanpak van die gemeten risico's bij die bepaalde patiënt. De soort gegevens die het assessment oplevert, wordt bepaald door de vragenlijsten die moeten worden beantwoord. De toepassing ervan berust op een aangenomen (deel)theorie, een assumptie, zoals: tbs-gestelden zijn vijandiger en minder empathisch dan anderen en die vijandigheid en egocentriciteit leiden tot delicten. Het afnemen van een vijandigheids- en empathievragenlijst ligt in dat geval voor de hand. Deze kunnen dan tevens worden afgenomen als evaluatie van de behandeling van die vijandigheid en egocentriciteit en van andere veranderingen in een patiënt. Voortdurende delictgevaarlijkheid wordt op deze wijze gespecificeerd, geobjectiveerd en (virtueel) gekwantificeerd over een verloop van tijd, met of zonder behandeling.

Risk management is dus te definiëren als een interventieplan op basis van risk assessment waarbij de interventies voortvloeien uit de uitkomsten van het assessment. Het effect van die interventies kan vervolgens afgemeten worden aan de veranderingen op de eerder afgenomen assessmentschalen. Het bijhouden van die veranderingen en de toegepaste therapievorm(en) daarbij wordt monitoring van het interventieprogramma genoemd. Deze monitoring kan heel

goed een onderdeel zijn van het vervolgen van een ambulante of klinische forensisch-psychiatrische behandeling in zijn geheel, waarbij de indicatie en duur van de verschillende interventies, plaats en kosten van verblijf en inspanningen ook worden gemeten. Op deze wijze is *benchmarking* van de verschillende behandelingen en behandelaars mogelijk: hun prestaties ten aanzien van een bepaalde problematiek kunnen onderling worden vergeleken, omdat dezelfde meetinstrumenten de parameters van die vergelijking zijn. Methodologisch zal echter wel aan een aantal voorwaarden moeten zijn voldaan, wil zo'n vergelijking ook tot keuzes kunnen leiden. Zo zullen behandelprogramma's goed gedefinieerd moeten zijn, de indicaties ervoor vastgelegd moeten zijn binnen een psychiatrisch (en niet alleen een gedragskundig) spectrum, en de settings waarin de interventies plaatsvinden als variabelen omschreven moeten zijn. Een duidelijk vastgelegd evaluatietraject en de bewaking van de integriteit ervan zijn onderdelen van monitoring. Op basis hiervan kunnen dan zowel korte- (acute) als langetermijnbeslissingen worden genomen die leiden tot een bijstelling van de interventies. De methodiek van het *case management* leent zich hier goed voor.

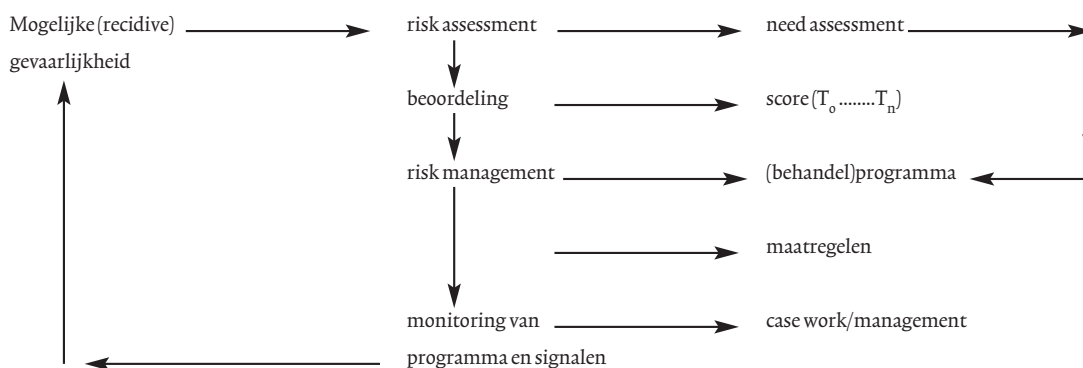
In figuur 1 wordt de risk-assessment-cyclus weergegeven die op deze manier ontstaat.

BESCHOUWING

Het combineren van het traditionele forensisch-psychiatrisch onderzoek met de methode van het risk assessment en risk management is veelbelovend. De combinatie lijkt een meerwaarde op te leveren, die zich vooral manifesteert in de vergroting van de onderlinge vergelijkbaarheid van verschillende patiënten ten opzichte van bepaalde uitkomsten (zoals bijvoorbeeld de effectiviteit van de behandeling van een bepaalde problematiek) en van de vergelijkbaarheid van dezelfde patiënt op verschillende momenten. Niet alleen gaat assessment uit van een impliciete, multicondionele theorie over crimineel gestoord gedrag, maar ook levert het een bijdrage aan verdere theorievorming: bepaalde hypothesen opgedaan uit demografisch en epidemiologisch onderzoek kunnen worden getoetst aan de uitkomsten van de assessment-schalen. Deze voortdurende evaluatie van die schalen ten opzichte van hun vermogen risico's te beoordelen doet de constructvaliditeit toenemen. Zo blijken de risk-assessmentschalen voor seksuele en niet-seksuele geweldsdelinquenten voor een gedeelte te moeten verschillen, wil de risicobeoordeling voor beide groepen optimaal zijn (Rice & Harris 1997; Hanson & Bussière 1998).

Hoewel in de Verenigde Staten en in Canada inmiddels veel ervaring bestaat met deze manier van werken, kunnen de gebruikte risicoschalen,

FIGUUR 1: Risk-assessment-cyclus



zoals de *Psychopathy Checklist-Revised* (PCL-R, Hare 1991), de *Violent Risk Appraisal Guide* (VRAG, Quinsey e.a. 1998) en de *Historical Clinical and Risk Management items-20 version 2* (HCR-20 version 2) (Webster e.a. 1997), niet zomaar worden overgenomen voor de Nederlandse forensische populatie. De wetsystemen in de eerdergenoemde landen zijn ten aanzien van de geestelijk gestoorde delinquenten geheel anders dan het onze, bijvoorbeeld omdat zij geen strafrechtelijke maatregel zoals de tbs kennen. Daar bestaat een tweedeling van delinquenten: geestelijk gestoord of niet gestoord, waarbij de eerste groep als psychotisch en ontoerekeningsvatbaar (*not criminal responsible of not guilty by reason of insanity*) verder wordt begeleid door de psychiatrische instituties. In de gevangnissen wordt het risk assessment vooral uitgevoerd bij vrijwillig gedetineerden die op die wijze verkorting van hun straf kunnen krijgen. In Nederland daarentegen is juist de combinatie van een breed gedefinieerde gestoordheid en gevaarlijkheid het criterium voor een forensisch-psychiatrische behandeling, zodat zowel patiënten met een persoonlijkheidsstoornis als die met een psychose in dezelfde forensisch-psychiatrische instituten worden behandeld. Daarnaast (Schoemaker & Van Zessen 1997) is er een groot aantal personen met een psychische stoornis in de verschillende penitentiaire inrichtingen bij wie eveneens risk assessment voor de hand ligt. De in Canada onderzochte en de in Nederland te onderzoeken populatie verschillen dus nogal.

Opmerkelijk is de Nederlandse forensisch-psychiatrische terughoudendheid bij het wetenschappelijk onderzoek naar fallometrie als diagnostisch instrument te midden van andere (Quinsey & Earls 1990) en naar de effectiviteit van medicamenteuze behandeling bij sommige zedelinquenten (Bradford 1990, 1996). Bij fallometrie wordt via de penisplethysmograaf (PPG) de bloeddorstrooming van de penis gemeten als maat voor seksuele opwinding bij het aanbieden van seksueel prikkelende afbeeldingen. Een significante score bij die meting in combina-

tie met psychopathologie zou 77% van de recidiveverkrachters goed voorspellen (Quinsey e.a. 1995). Gecontroleerd klinisch onderzoek behoort tot de mogelijkheden. In ieder geval zou hier een discussie in ethische, juridische, methodologische, diagnostische en therapeutische zin kunnen worden gevoerd over de toepassing ervan. Wegens het ontbreken van een wetenschappelijk gefundeerd oordeel zijn beide methodes de speelbal van politieke en ideologische sentimenten.

De betekenis van het risk assessment voor de Nederlandse forensische psychiatrie moet nog blijken. De waarde ervan zal pas bekend zijn nadat na een aantal jaren de recidivepercentages zijn berekend van de populatie die na een gestandaardiseerde risicobeoordeling is ontslagen. Gezien de mogelijke grote voordelen heeft deze methode nu al recht op een plaats binnen het forensisch diagnostisch instrumentarium.

LITERATUUR

- Blazer, D.G., & Hays, J.C. (1998). *An introduction to clinical research in psychiatry*. Oxford: Oxford University Press.
- Bonta, J., Hanson K., & Law M. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders. A meta-analysis. *Psychological Bulletin*, 123, 123-142.
- Bradford, J.M.W. (1990). The antiandrogen and hormonal treatment of sex offenders. In W.L. Marshall, D.R. Laws & H.E. Barbaree (red.), *Handbook of sexual assault. Issues, theories and treatment of the offender* (pp. 297-310). New York: Plenum Press.
- Bradford, J.M.W. (1996). The role of serotonin in the future of forensic psychiatry. *Bulletin of the American Academy of Psychiatry and the Law*, 24, 57-72.
- Brand, E.F.J.M., Diks, G.J.M., van Emmerik, J.L., e.a. (1998). *Diagnostiek en onderzoek. De ontwikkeling van een gestructureerd diagnostisch instrumentarium voor klinische en researchdoeleinden*. Utrecht: F.S. Meijersinstituut, monitoring & research.
- Campbell, J.C. (red.) (1995). *Assessing dangerousness. Violence by sexual offenders, batterers and child abusers*. California, USA: IVPS, Sage Publications.
- Denzin, N.K., & Lincoln, Y.S. (red.) (1994). *Handbook of qualitative research*. Thousand Oaks: Sage Publications Inc.
- Gardner, W., Lidz, C.W., Mulrey, E.P., e.a. (1996). *Clinical versus actu-*

- arial predictions of violence in patients with mental illnesses. *Journal of Consulting and Clinical Psychology*, 64, 602-609.
- Grubin, D. (1999). Actuarial and clinical assessment of risk in sex offenders. *Journal of Interpersonal Violence*, 14, 331-343.
- Haen Marshall, I. (1998). De predictie van geweldscriminaliteit. *Tijdschrift voor Criminologie* 1998/1, 2-20.
- Hanson, R.K., & Bussière, M.T. (1998). Predicting relapse. A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348-362.
- Hare, R.D. (1991). *Manual for the Hare Psychopathy Checklist Revised*. Toronto: Multihealth Systems.
- Hildebrand, M., & de Ruiter, C. (1999). Classificatie en diagnostiek in de forensische psychiatrie. In C. de Ruiter & M. Hildebrand (red.), *Behandlungsstrategieën bij forensische psychiatrische patiënten*. Houten/Diemen: Bohn Stafleu Van Loghum.
- Litwack, T.R., Kirschnner, S.M., & Wack, R.C. (1993). The assessment of dangerousness and predictions of violence: Recent research and future prospects. *Psychiatric Quarterly*, 64, 245-273.
- Marle, H.J.C. van (1986). Het voorspellen van gevaarlijk gedrag. *Tijdschrift voor Psychiatrie*, 28, B2, 1-9.
- Marle, H.J.C. van (1992). Voorspellen van gevaarlijkheid. Consensus gewenst. *Tijdschrift voor Psychiatrie*, 34, 281-290.
- Marle, H.J.C. van (1999a). Het strafrechtelijk gedragskundig onderzoek. In B.C.M. Raes & F.A.M. Bakker (red.), *De psychiatrie in het Nederlands recht* (pp. 75-94). Deventer: Gouda Quint.
- Marle, H.J.C. van (1999b). TBS op maat. Een overzicht van de discussie. *Justitiële verkenningen* 4/99 (pp. 40-53). Gouda Quint: WODC Justitie.
- McGuire, J. (1995). *What works; Reducing offending*. Chichester: Wiley.
- Monahan, J., & H.J. Steadman (1994). Toward a rejuvenation of risk assessment research. In J. Monahan & H.J. Steadman (red.), *Violence and mental disorder. Developments in risk assessment* (pp. 1-17). Chicago: University of Chicago Press.
- Mooij, A.W.M. (1998). *Psychiatrie, recht en de menselijke maat. Over verantwoordelijkheid*. Amsterdam: Boom.
- Quinsey, V.L., & Earls, C.M. (1990). The modification of sexual preferences. In W.L. Marshall, D.R. Laws & H.E. Barbaree (red.), *Handbook of sexual assault. Issues, theories and treatment of the offender* (pp. 279-295). New York: Plenum Press.
- Quinsey, V.L., Lalumière, M.L., Rice, M.E., e.a. (1995). Predicting sexual offenses. In J.C. Campbell (red.), *Assessing dangerousness. Violence by sexual offenders, batterers and child abusers* (pp. 114-137). Thousand Oaks: Sage Publications.
- Quinsey, V.L., Harris, G.T., Rice, M.E., e.a. (1998). *Violent offenders appraising and managing risk*. Washington DC: American Psychological Association.
- Raine, A. (1993). *The psychopathology of crime. Criminal behavior as a clinical disorder*. San Diego, Cal.: Academic Press, Inc.
- Rice, M.E., & Harris, G.T. (1997). Cross-validation and extension of the violent risk appraisal guide for child molesters and rapists. *Law and Human Behavior*, 21, 231-241.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118, 183-192.
- Schoemaker, C., & van Zessen, G. (1997). *Psychische stoornissen bij gedetineerden*. Utrecht: Trimbos-instituut.
- Steadman, H.J., Monahan, J., Appelbaum, P.S., e.a. (1994). Designing a new generation of risk assessment research. In J. Monahan & H.J. Steadman (red.), *Violence and mental disorder. Developments in risk assessment* (pp. 297-318). Chicago: University of Chicago Press.
- Verhagen, M.F.M., & Philipse, M.G.W. (1995). Het voorspellen van het risico van delictherhaling in de TBS. *Tijdschrift voor Psychiatrie*, 37, 537-552.
- Webster, C.D., Ben-Aron, M.H., & Hucker, S.J. (1985). *Dangerousness. Probability & prediction, psychiatry & public policy*. Cambridge, London, New York: Cambridge University Press.
- Webster, C.D., Douglas, K.S., Eaves, D., e.a. (1997). *HCR-20. Assessing risk for violence, version 2*. British Columbia: Simon Fraser University.

AUTEUR

H.J.C. VAN MARLE is hoogleraar forensische psychiatrie aan de Katholieke Universiteit Nijmegen en psychiatrisch adviseur van het ministerie van Justitie; tevens is hij psychoanalytisch psychotherapeut bij de H. van der Hoevenkliniek in Utrecht.
Correspondentieadres: prof. dr. H.J.C. van Marle, Katholieke Universiteit Nijmegen, Thomas van Aquinostraat 6, 6525 GD Nijmegen.

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 10-4-2000.

SUMMARY

From danger to risk: Managing instead of warning. Risk assessment in forensic psychiatry – H.J.C. van Marle –

Assessing risk before leave and termination of forensic treatment is very important because of the big consequences for society and the patient himself. The usual clinical examination by a psychiatrist or a psychologist depends very much on the setting, the experience and the expertise of the examiner. The introduction of a more objective standard for decision making is recommended for the whole process of reporting to the courts and other decision makers. Experiences with risk assessment scales, like those developed in Canada, are hopeful but validation is needed for the Dutch situation. This more objective assessment though is not able to replace the examination for the degree of criminal responsibility as only the hermeneutic method gives insight in the specific combination of offender and offence. Also during the forensic psychiatric treatment individual examination remains essential for the planning of the therapeutic process. The expectation is that the combination of the individual and the assessment approach will make the treatment more specific and more effective concerning the risks.

The implementation of the risk assessment/risk management cycle within the forensic field brings many advantages by following and evaluating the forensic patients intensively. It is well known that some of them can be a danger for society for many years after a completed treatment.

[TIJDSCHRIFT VOOR PSYCHIATRIE 42 (2001) 3, 151-161]

KEYWORDS forensic psychiatric report, hermeneutics, risk assessment, risk management