

Eric Blaauw*, Stefan Bogaerts** en Marinus Spreen***

Risicotaxatie in de Nederlandse rechtspraak: op naar een best practice

Recent is veel discussie ontstaan over de gangbare methode voor risicotaxatie in de Nederlandse forensische diagnostiek. In dit artikel bespreken wij de voor- en nadelen van de verschillende methoden vanuit zowel de wetenschap als de rechtspraak en beargumenteren wij waarom het gestructureerd klinisch oordeel, naar onze mening, de beste methode is voor risicotaxaties ten behoeve van de rechtbank. Daarbij bespreken wij ook de vraag wie het best geplaatst is om een risicotaxatie te verrichten, of een rechter ook zelf risicotaxaties kan verrichten, en of een pro-justitia-rapporteur een inschatting van het recidiverisico behoort te maken wanneer er geen stoornis wordt vastgesteld. Ook beperkingen van risicotaxatie komen aan de orde in dit artikel, waardoor het mogelijk is een aanzet te geven tot een best practice van risicotaxatie in Nederland.

Inleiding

In 2017 en 2018 verschenen verschillende berichten in de media en vakbladen over de belangrijke rol van risicotaxatie in de forensische keten. Zo bleek de verdachte in de zaak-*Anne Faber* reeds eerder veroordeeld te zijn voor twee soortgelijke delicten en werd in de publieke opinie de vraag opgeworpen waarom er toentertijd geen terbeschikkingstelling met dwangverpleging was opgelegd. In de berichten werd het gebruik van actuariële risicotaxatie-instrumenten, ook wel algoritmische risicotaxatie-instrumenten genoemd, naar voren geschoven als de meest accurate methode om risico's van verdachten en forensisch psychiatrische patiënten in te schatten. De Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen concludeerde dat het recidief delict-risico onder zedendelinquenten in Nederland niet accuraat wordt ingeschat, omdat in plaats van actuariële of algoritmische instrumenten, gebruik wordt gemaakt van instrumenten die zich baseren op het Gestructureerd Klinisch Oordeel (GKO). In het *Tijdschrift voor Psychiatrie* trokken twee onderzoekers van de Universiteit van Amsterdam dit standpunt verder door naar alle risicotaxaties door te stellen: 'In Nederland wordt het overweldigende wetenschappelijke bewijs dat actuariële risicotaxatie beter presteert dan klinische risicotaxatie grotendeels genegeerd in jeugd- en forensische zorg.'¹

In dit artikel willen wij een bijdrage leveren aan deze discussie door de voor- en nadelen van de verschillende risicotaxatie-benaderingen te bespreken en tevens de vraag te verkennen in hoeverre rechters zelf risicotaxaties bij rechtszaken en verlengingsadviezen tbs kunnen gaan verrichten. Daarbij stellen wij tevens de vraag aan de orde of pro-justitia-rapporteurs ook een inschatting van het recidiverisico behoren te maken, wanneer er geen psychische stoornis wordt vastgesteld. Alvorens op deze vragen in te gaan, geven wij eerst een kort overzicht van hoe risicotaxatie in de Nederlandse rechtspraak wordt ingezet en een beknopte geschiedenis van de risicotaxatie.

Risicotaxatie in de Nederlandse rechtspraak

In Nederland worden risicotaxaties uitgevoerd ten behoeve van de rechtspraak en ten behoeve van de forensische behandeling om het recidiverisico te schatten en de behandelvoortgang te meten. In de rechtspraak verrichten pro-justitia-rapporteurs risicotaxaties bij slechts een fractie van de verdachten die nog voor de rechter moeten komen. Hierbij is de risicotaxatie dan vaak ondergeschikt aan de vraag in hoeverre een psychische stoornis van invloed was op het tenlastegelegde delict. Daarnaast worden risicotaxaties door pro-justitia-rapporteurs verricht voor de rechtbank bij alle terbeschikkinggestelden in het kader van de vier-jaarrapportages en bij onderzoeken of de maatregel van tbs onvoorwaardelijk kan worden beëindigd. Ten slotte worden risicotaxaties voor de rechtbank verricht door de reclassering die bij het voorlichtingsrapport gebruik kan maken van de Recidive Inschattingsschalen (RISc) en de recent toegevoegde Oxford Risk of Recidivism tool (OxRec). Het pro-justitia-onderzoek heeft als doel het inzichtelijk maken van risicofactoren die samenhangen met het delict om daarmee het recidiverisico in te schatten en vervolgens aanbevelingen te formuleren omtrent interventies die het eventuele recidivegevaar kunnen verminderen. Risicotaxatie in een behandelsetting kent een ander doel, namelijk het periodiek (gemiddeld eens per jaar) inschatten van de verandering van het recidiverisico en het meten van de behandelvoortgang. Hierbij is het GKO essentieel: naast het systematisch en gestructureerd scoren van risico- en beschermende factoren conform de handleiding (scoringsintegriteit), dient de beoordelaar alle informatie te integreren om te komen tot een inschatting van het recidiverisico. Ten behoeve van de forensische behandeling en het monitoren van de voortgang hierin (Routine Outcome Monitoring: ROM), zijn de forensisch psychiatrische centra sinds 2004 verplicht om jaarlijks een risicotaxatie te verrichten bij terbeschikkinggestelden (aan de hand van de HKT-R of HCR-20^{V3}, met daarnaast SVR-20 en PCL-R bij verlofaanvragen) en is het sinds 2015 ook binnen andere settings voor forensische

* Dr. Eric Blaauw, Hanzehogeschool Groningen en Verslavingszorg Noord Nederland.

** Prof. dr. Stefan Bogaerts, Tilburg University en Fivoor Wetenschap en Behandelinovatie.

*** Dr. Marinus Spreen, NHL Stenden Hogeschool en FPC Dr. S. van Mesdag.

1. Van der Put & Assink 2018, p. 510.

zorg verplicht een risicotaxatie te verrichten bij justitiabelen die zorg krijgen vanuit een (voorwaardelijke) straf of maatregel of andere strafrechtelijke titel (HKT-R voor klinische zorg en FARE voor ambulante zorg).

In dit artikel beperken we ons tot risicotaxaties ten behoeve van de rechtbank, omdat verschillende contexten met betrekking tot risicotaxatie verschillende vraagstukken en uitvoeringsproblemen kennen.

Geschiedenis van risicotaxatie

Oorspronkelijk werd het ongestructureerde klinische oordeel gebruikt voor risico-inschattingen ten behoeve van de rechtbank. Daarbij maakte een gedragsdeskundige op basis van eigen expertise en klinische indrukken een inschatting van het recidiverisico van een verdachte. De gedachte was daarbij dat gedragsdeskundigen, vanuit hun theoretische kennis en ervaring, inzicht en kennis hadden van associaties die bestaan tussen psychische stoornissen en het recidiverisico. Omdat in veel gevallen hierbij niet werd vermeld op welke empirische gegevens het oordeel was gebaseerd, en de klinische besluitvorming weinig transparant verliep en bijgevolg niet replicerbaar was, werden vooral in Noord-Amerika risicotaxatie-instrumenten ontwikkeld die meer op wetenschappelijk onderzoek waren gebaseerd, minder variatie kenden tussen beoordelaars en een hoge mate van transparantie en replicerbaarheid lieten zien. Hiertoe werden algoritmische instrumenten ontworpen met risicofactoren die gerelateerd waren aan recidive en afkomstig waren van grote databestanden waarin werd gezocht naar statistisch discriminerende kenmerken tussen recidivisten en niet-recidivisten.

Een algoritmische risicotaxatie houdt in dat de onderzoeker gebruikmaakt van een beperkt aantal variabelen waarvan empirisch is aangetoond dat ze een verhoogde kans op recidive meten. Vervolgens worden de variabelen gecodeerd en via een vaststaand algoritme (meestal via het optellen van de scores) wordt een oordeel van het recidiverisico verkregen. Kenmerkend voor deze risicotaxatie-benadering is dat het risico wordt bepaald door het algoritme en niet door de onderzoeker, wat tevens betekent dat er geen sprake is van een GKO. Een ander kenmerk van de algoritmische benadering is dat de variabelen die moeten worden gescoord objectief zijn vast te stellen: het gaat om min of meer 'administratieve variabelen', zoals het aantal eerdere veroordelingen, leeftijd ten tijde van het eerste delict, soort delict, bekend of onbekend slachtoffer, enzovoort. Met het oog op het bevorderen van transparantie van de onderliggende besluitvormingsprocessen van de risicotaxatie, lijkt de conclusie van zowel de Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen als van Van der Put en Assink op het eerste gezicht dus niet onredelijk.

Daar waar de internationale literatuur in eerste instantie veel nadruk legde op het ontwikkelen van algoritmische risicotaxatie-instrumenten, werd deze methode ook als

onbevredigend ervaren. Met de algoritmische benadering werd namelijk geen rekening gehouden met individuele kenmerken van de verdachte of patiënt. Bovendien zijn de eerder vermelde 'administratieve kenmerken' waarop de algoritmische methode zich baseert onveranderbaar. Anders gezegd: wanneer een verdachte met een hoog risico werd veroordeeld tot een (deels) voorwaardelijke straf of maatregel, dan bleef het risico op basis van een algoritmische methode hierna even hoog, ondanks een eventuele geslaagde behandeling. Ook mag worden gesteld dat beslisregel gebaseerde algoritmen niet zaligmakend zijn; valspositieve en valsnegatieve voorspellingen zijn niet uit te sluiten en algoritmes zijn nooit volledig duidelijk en transparant. Vanwege deze beperkingen werden de huidige gepersonaliseerde GKO-instrumenten ontwikkeld.

Een GKO-instrument bestaat uit een aantal risicofactoren die meestal zijn onderverdeeld in drie tijdsdomeinen.² Het eerste (historische) tijdsdomein bestaat uit factoren die in principe niet kunnen veranderen vanwege hun historisch karakter, zoals eerdere veroordelingen en leeftijd ten tijde van het eerste delict. De vergelijking met de algoritmische instrumenten ligt op het eerste gezicht voor de hand. Echter, bij sommige historische variabelen is sprake van een subjectieve inschatting door de gedragsdeskundige. Dit zijn variabelen als inschatting van de ernst van de problematiek, problematische woongeschiedenis en arbeidsverleden. Het tweede (klinische) tijdsdomein bevat variabelen die risicovolle gedragingen van de afgelopen periode (meestal de afgelopen 12 maanden) in kaart brengen, zoals de mate van impulsiviteit, empathie, vijandigheid en psychotische episodes. Deze variabelen zijn dynamisch, omdat ze per periode kunnen veranderen door omstandigheden, zoals gevangenisstraf en/of behandeling. Ten slotte behelst het derde domein variabelen die een inschatting van de toekomst vereisen. Het gaat er dan om wat de verdachte of patiënt op allerlei levensgebieden (nog) nodig heeft om zonder delicten door het leven te gaan, zoals het hebben van een woning en een ondersteunend netwerk. Van de individuele variabelen uit de GKO-instrumenten zijn, net als de algoritmische instrumenten, vanuit wetenschappelijk onderzoek associaties met recidive vastgesteld. Een GKO-risicotaxatie bestaat dan uit het invullen van de diverse domeinen, zodat een goed beeld wordt verkregen van de risico's van de verdachte en/of patiënt. Vervolgens dient de gedragsdeskundige andere omstandigheden die niet zijn gemeten door het instrument, in het eindoordeel mee te wegen. Aanbevolen voor het werken met een GKO-instrument is dat minstens twee getrainde professionals onafhankelijk van elkaar het instrument invullen en in een consensusbespreking tot een onderbouwd gepersonaliseerd risico-oordeel komen. Men kan dus stellen dat de GKO-werkwijze een combinatie is van de ongestructureerde en de algoritmische werkwijze, waarbij gebruik wordt gemaakt van zowel statistische gegevens als van klinische inschattingen.

2. Bogaerts, Spreen, Ter Horst & Gerlisma 2018.

Algoritmisch instrument of gestructureerd klinisch oordeel? De wetenschap

Al decennia terug beargumenteerde Meehl³ in zijn boek *Clinical vs. Statistical Prediction: A Theoretical Analysis and A Review of the Evidence* dat de algoritmische methode beter voorspelt dan de klinische methode. Tegelijkertijd stelde Meehl ook dat de algoritmische voorspelling hoogst uitzonderlijk mag worden bijgesteld door de clinicus, wanneer substantiële informatie niet is opgenomen in het algoritmische voorspellingsmodel en wanneer deze informatie de besluitvorming drastisch zou wijzigen. Hieruit volgde dus de ontwikkeling van GKO-instrumenten.

Ook in voorbije decennia speelde de vraag of de algoritmische risicotaxatie dan wel het GKO leidt tot betere voorspellingen van het recidiverisico. Aanvankelijk lieten diverse meta-analyses onder delict-plegers,⁴ delinquenten met al dan niet een psychische stoornis⁵ en zedendelinquenten⁶ zien dat de algoritmische methode tot een betere voorspelling leidde dan het klinisch oordeel. Hierbij werd bij de meta-analyses vooral het aloude ongestructureerde klinisch oordeel afgezet tegen algoritmische methoden en werd niet of onvoldoende de vergelijking gemaakt met het wetenschappelijk gefundeerde GKO. Uit recentere meta-analyses⁷ blijkt dat de gemiddelde predictieve validiteit van GKO-oordelen voor het voorspellen van recidive met een geweldsdelict of enig delict, zeker niet slechter is dan dat van algoritmische instrumenten. Er is dus met betrekking tot het voorspellen van deze typen delicten zeker geen overweldigend wetenschappelijk bewijs dat de algoritmische risicotaxatie beter zou presteren dan het GKO, zoals Van der Put en Assink stellen.

Omtrent de voorspelling van recidive met een zedendelict concluderen Hanson en Morton-Bourgon⁸ op basis van meta-analyses dat actuariële instrumenten een betere voorspelling opleveren dan het GKO bij deze groep delinquenten. De vergelijking gaat echter een beetje mank, omdat wat GKO-instrumenten betreft, alleen werd gekeken naar de SVR-20 wat een instrument is zonder recidivecijfers bij verschillende risicocategorieën, waardoor de SVR-20 dus minder richtinggevend is dan andere GKO-instrumenten. De Static-99R (en ook de Stable-2007) heeft het voordeel dat gebruik kan worden gemaakt van recidivecijfers bij de categorieën van gevaarinschattingen. Met het oog op transparantie van inschattingen hebben de Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen en Van der Put en Assink dus zeker een punt wat risicotaxatie bij zedendelinquenten betreft. Maar de predictieve validiteit van de Static-99R is slechts bescheiden en de normtabellen zijn afkomstig uit Noord-Amerika, wat de vraag oproept hoe de recidivecijfers zich verhouden tot Nederland. Bovendien blijkt uit eerder onderzoek⁹ dat de SVR-20 in Nederland betere voorspel-

lingen opleverde dan de vroegere versie van de Static en beter presteerde dan was weergegeven in de meta-analyse van Hanson en Morton-Bourgon. Met andere woorden: het overweldigend wetenschappelijk bewijs waarvan Van der Put en Assink spreken is wat teleurstellend.

Bij de meta-analyses moeten enkele kritische kanttekeningen worden gemaakt. Allereerst bestaan de populaties in de meta-analyses uit een variëteit van juridische maatregelen, van alleen gevangenisstraf, ambulante maatregelen tot gedwongen psychiatrische behandelingen. Hierdoor zijn instrumenten niet zonder meer bruikbaar in Nederland door verschillen in dadergroepen, verschillen in rechtssystemen en verschillen in strafbaarstelling van bepaalde delicten. Ook laten de meta-analyses duidelijk zien dat ook algoritmische methoden slechts een beperkte voorspelling geven van het recidiverisico – oftewel veel van de variatie blijft onverklaard. Daarbij geven vooral hoge en lage scores redelijk accurate voorspellingen, omdat relatief weinig onderzochten hoog scoren en het recidivepercentage in de groep hoogscoorders doorgaans hoog is en omdat veel onderzochten laag scoren en het recidivepercentage in de groep laagscorders doorgaans laag is.¹⁰ Ook kunnen enkele kritische kanttekeningen worden geplaatst bij de totstandkoming van algoritmische instrumenten. Bij risicotaxatie-instrumenten geldt dat het instrument als goed wordt gekwalificeerd als willekeurige recidivisten in 80 procent van de gevallen hoger scoren op het instrument dan willekeurige niet-recidivisten. Uit de meta-analyses is af te leiden dat geen van de algoritmische instrumenten momenteel aan deze kwaliteitseis voldoet. Ook is sociale ongelijkheid ingebakken in de modellen en kunnen variabelen die geen oorzakelijk verband hebben met recidive (zoals etniciteit of muziekvoorkeur), terecht komen in een algoritmisch instrument en kunnen daardoor leiden tot ethische vraagstukken, zoals terecht wordt geconstateerd in het tijdschrift *Strafblad*.¹¹ Alle berichtgevingen, studies en speculaties ten spijt, zijn derhalve de nodige kanttekeningen te plaatsen bij de algoritmische methode van risicotaxatie.

Algoritmisch instrument of gestructureerd klinisch oordeel? De rechtspraak

In eerste instantie lijkt een zuiver algoritmisch instrument wenselijk voor de rechtspraak omdat verdachten hierbij worden getaxeerd door middel van een aantal criminogene variabelen (zoals leeftijd, geslacht, aantal eerdere delicten, enz.) die in principe foutloos kunnen worden ingevuld en in theorie door een naïeve codeur kunnen worden bepaald. Daarbij bepaalt het algoritme en niet de codeur de hoogte van het risico en is er de veronderstelling dat er sprake is van een hoge mate van validiteit en betrouwbaarheid en van transparantie. Nadeel is echter dat de persoonlijke omstandigheden van de verdachte niet kunnen worden meegenomen om het risiconiveau te nuanceren. Daardoor is het mogelijk dat

3. Meehl 1954.

4. Grove e.a. 2000.

5. Bonta e.a. 1998.

6. Hanson & Bussière 1998.

7. Fazel e.a. 2012; Guy 2008; Singh, Grann & Fazel 2011.

8. Hanson & Morton-Bourgon 2004, 2009.

9. De Vogel e.a. 2004.

10. Zie bijvoorbeeld tabel 10 van de handleiding HKT-R, Spreen e.a. 2014.

11. Van Wingerden 2018.

een verdachte die geen enkele justitiële voorgeschiedenis heeft en twee kinderen seksueel heeft misbruikt, een laag risico krijgt volgens het algoritme, terwijl de verdachte bijvoorbeeld geobsedeerd is door het misbruiken van kinderen, getuige zijn uitspraken en de kinderporno op zijn computer. Volgens de algoritmische benadering dient men dan uit te gaan van het vastgestelde risico. Het moge duidelijk zijn dat het in dit geval aannemelijk is het risico als hoger te beoordelen, iets wat met behulp van een GKO-instrument mogelijk is.

Te stellen is dat het gebruik van algoritmische risicotaxatie niet goed strookt met de werkwijzen van rechters. Een strafrechter kent bij de beoordeling van delicten een voorgeschreven reikwijdte van strafoplegging in het wetboek en heeft daarbij een grote beslissingsbevoegdheid, waarbij vanuit een utilitaristische achtergrond vooral kan worden gewerkt aan het voorkomen van schade aan de samenleving en aan het resocialiseren van de dader. Ook kan het vanuit een overweging van retributie zinvol zijn om alternatieve of voorwaardelijke straffen op te leggen. Indien het wetboek algoritmisch zou zijn opgebouwd, dan zou dit betekenen dat bij een delict een vaststaande straf bestaat en dat de rol van de rechter zich beperkt tot het vaststellen van de schuld. Dit is natuurlijk niet de praktijk, omdat het werk van de rechter er ook uit bestaat om de persoonlijke omstandigheden van de verdachte mee te wegen in de strafmaat. Deze keuzevrijheid zien we terug in de bevindingen van onderzoek onder rechters¹² dat liet zien dat de statistische kans op recidive geen leidende straftoemingsfactor is bij de rechters en dat vooral het feitelijke gedrag van het individu (samen met de tijd die de verdachte in voorlopige hechtenis heeft doorgebracht) bepalend is voor de straftoemeting. Vanuit de analogie met de werkwijze van rechters ligt dan het GKO meer voor de hand, omdat het bij deze methodiek mogelijk is de persoonlijke omstandigheden in te schatten en mee te wegen.

Problemen bij de inschatting van het recidiverisico

In 2010 concludeerden Harte en Breukink vanuit een beschouwing van risicotaxatie-instrumenten die in de Nederlandse praktijk worden gebruikt, dat de predictieve validiteit van de instrumenten nog te wensen overlaat. Daarbij speelt tevens een rol dat verschillende risicotaxatie-instrumenten verschillende zaken voorspellen. Zo beoogt de RISC een indicatie te geven van het risico op recidive met een algemeen delict, richt de HKT-R zich op het risico op een nieuw geweldsdelict in een periode van vijf jaar en adresseert de Static-99R een nieuwe aanklacht of veroordeling voor een seksueel delict (of een geweldsdelict) na vijf of tien jaar. Het is daarom erg van belang om het juiste instrument te kiezen voor wat men wil voorspellen. Ook is het van belang om het juiste instrument bij de juiste dadergroep te kiezen, omdat instrumenten voor specifieke dadergroepen doorgaans betere voorspellingen opleveren dan meer algemene instrumenten,¹³ omdat zij geijkt zijn op deze specifieke

groepen. Zo is het bijvoorbeeld verstandig om bij vrouwelijke daders gebruik te maken van een toegevoegd instrument dat is geijkt op vrouwelijke daders (FAM¹⁴). Inmiddels zijn meerdere instrumenten doorontwikkeld op specifieke dadergroepen en voor specifieke voorspellingen, maar is er zeker nog geen reden om te stellen dat de predictieve validiteit zo goed is dat rechters zich hierover geen zorgen hoeven te maken.

Een bijzonder probleem is dat de meeste algoritmische en/of GKO-instrumenten alleen zijn getest op hun predictieve validiteit in groepen delictplegers met een psychische stoornis. Dit komt omdat deze instrumenten zijn ontwikkeld vanuit een behandelingsperspectief. Vanuit behandelingsperspectief wil men het recidiverisico en de delictgerelateerde risicofactoren kennen, zodat men kan inschatten wat de behandelintensiteit dient te worden, welke specifieke dynamische risicofactoren dienen te worden behandeld en welke behandeling het best aansluit bij de leefwereld en mogelijkheden van de patiënt. Dit behandelingsperspectief wordt vooral ingevuld vanuit het Risk Need Responsivity- (RNR)-model.¹⁵ De meeste instrumenten zijn dan ook gevalideerd op recidive na behandeling. En hier wringt de schoen op twee verschillende punten. Het eerste knelpunt is dat bij pro-justitia-rapportages nog niet duidelijk is of er sprake zal zijn van vrijlating, gevangenisstraf of een behandeling vanuit voorwaarden. Het tweede knelpunt is dat bij validering van een instrument waarin de behandeling of straf is meegenomen (het gaat immers om recidive na deze behandeling of straf) vooral ook het effect van de straf of behandeling wordt verdisconteerd in het instrument. Deze knelpunten zijn alleen op te lossen met instrumenten die zijn gevalideerd op de pro-justitia-fase zonder dat er sprake is van een selectie voor behandeling. Helaas zijn die instrumenten er (nog) niet, zodat voorzichtigheid geboden is voor rechters bij de interpretatie van de uitkomsten van de huidige ingezette instrumenten.

Helaas is het niet mogelijk om voorspellingen met 100 procent zekerheid te maken, maar moet een betrouwbaarheidsinterval in acht worden genomen. Het grootste probleem voor rechters is echter dat de recidivecijfers voor de verschillende risicocategorieën geen duidelijke handvatten bieden over hoe hiermee moet worden omgegaan. Zo is het percentage recidivisten met een geweldsdelict vanuit de HKT-R 16 procent bij een lage score, 26 procent bij een matige score en 52 procent bij een hoge score (zie handleiding HKT-R). Vrijlating van iedereen met een lage score, zou dan bij 16 procent van deze groep delictplegers leiden tot recidive met een geweldsdelict. Evenzo zou het vasthouden van iedereen met een hoge score dan in 48 procent van deze gevallen onterecht zijn, omdat er geen sprake van recidive is. In beide gevallen is hiermee moeilijk te werken. Met risicotaxatie-instrumenten voor zeden is het al niet veel beter gesteld. Daarbij laat de handleiding van de Static-99R zien dat het percentage nieuwe aanklachten of veroordelingen voor een nieuw seksueel delict bij personen die zijn voorgeselecteerd voor behandeling van seksuele recidive (let wel: dit is dus slechts een selectie van alle

12. Zie Van Wingerden, Moerings & Van Wilsem 2011.

13. Singh, Grann & Fazel 2011.

14. De Vogel e.a. 2012.

15. Andrews, Bonta & Hoge 1990.

verdachten van een zedendelict) ongeveer 40 procent is na vijf jaar bij de hoogste score op de Static-99R en dat de kans op een nieuw geweldsdelict binnen een vergelijkbare groep ongeveer 56 procent is. Wanneer ook de klinische items van de Stable-2007 worden beoordeeld, lijken de voorspellingen wat beter te worden, maar zeker niet voldoende te zijn voor zekere besluitvorming.

Wie kan een gestructureerd klinisch oordeel geven?

Een belangrijke vraag is wat nodig is om een GKO te kunnen geven. Duidelijk is dat dit in theorie niet kan worden gedaan door een naïeve codeur, omdat meer informatie nodig is over het functioneren van de verdachte dan alleen naar voren komt uit voorgedeede items en omdat de taxateur een beargumenteerde inschatting moet geven over met name niet-historische kenmerken. Bestudering van handleidingen laat zien dat deskundigheid nodig is op het gebied van individuele psychodiagnostiek en psychische stoornissen en kennis van wetenschappelijke literatuur over geweld (HCR-20^{V3}) of dat de taxateur een diagnosticus, onderzoeker of hoofdbehandelaar moet zijn (HKT-R). Hierbij is het van belang om te weten dat de psychiatrische diagnose op zichzelf geen tot weinig verband heeft met het recidiverisico en dat het bij het recidiverisico vooral gaat om de onderliggende symptoomclusters of karakterclusters. Bij inspectie van de items van de meest gebruikte risicotaxatie-instrumenten in Nederland, zoals de HKT-R en de HCR-20^{V3}, blijkt dat de psychiatrische diagnose zelfs niet wordt meegewogen. De indicatoren waarvan bekend is dat zij enige relatie hebben met recidive, bestaan wat betreft het klinische gedeelte voornamelijk uit risico verhogende gedragingen en kenmerken, zoals de mate van impulsief gedrag, verslaving, gebrek aan empathie en gebrek aan sociale en copingvaardigheden. Deze kenmerken en gedragingen kunnen bij ieder mens in meerdere of mindere mate voorkomen en limiteren zich niet tot een psychiatrische diagnose. Sterker nog: een psychiatrische diagnose is niet noodzakelijk voor het doen van een risicotaxatie. Stel dat een verdachte geen diagnose heeft maar hoog scoort op de klinische schaal van een risicotaxatie-instrument (bijvoorbeeld lid van een jeugdbende, heel impulsief en onder invloed van alcohol agressief), dan is dit al snel voldoende om een hoog recidiverisico te veronderstellen. Daarentegen hoeft men bij een 17-jarige, alleenstaande *first offender* die op het moment van het delict eenmalig een psychotische episode had als gevolg van een uit de hand gelopen experiment met lsd en als gevolg daarvan een zedendelict beging jegens een onbekende jongen, nauwelijks bang te zijn voor recidive, terwijl er wel een psychiatrische stoornis was (en uit de Static-99R een hoog recidiverisico volgt).

Men kan stellen dat een pro-justitia-rapporteur als deskundige op het gebied van individuele psychodiagnostiek en psychische stoornissen bij uitstek geschikt is voor het geven van een GKO omtrent het recidiverisico. Toch is deze stelling niet zonder meer juist, omdat een deskundige op het gebied van psychodiagnostiek niet automatisch tevens deskundig is op het gebied van risicotaxatie. Voor een gekwalificeerde pro-justitia-rapporteur is jarenlange scholing en deskundigheidsbevordering vereist. Voor risicotaxatie wordt vaak volstaan met

aanvullende training en het bieden van informatie omtrent risicofactoren en recidiverisico's van verschillende typen geweld (waaronder zeden), maar is het maar de vraag of dit afdoende is. Bovendien is de stelling niet juist vanuit de voorgenoemde wetenschap dat het niet zozeer de vastgestelde stoornissen zijn die samenhangen met een verhoogd recidiverisico, maar eerder de symptomen, karaktertrekken en situationele omstandigheden en dat vooral gedegen onderzoek van deze symptomen, karaktertrekken en situationele omstandigheden van belang is voor de inschatting van het recidiverisico. Het gevaar bestaat dat een pro-justitia-rapporteur bij de vaststelling van een stoornis, zich er onvoldoende van bewust is dat dit niet automatisch samenhangt met een verhoogd recidiverisico of dat de afwezigheid van een stoornis niet automatisch betekent dat er geen recidiverisico aanwezig is. De rol van de pro-justitia-rapporteur is primair het stellen van een diagnose en het onderzoeken van de relatie van de stoornis met het tenlastegelegde delict. Voor de rechter is het hieruit resulterende advies belangrijk voor de beantwoording van de vraag of het tenlastegelegde feit volledig is toe te rekenen aan de verdachte en of het opleggen van een justitiële maatregel geïndiceerd is ter vermindering van de kans op recidive of ter beveiliging van de maatschappij. Daarnaast is de diagnose voor de rechter meer zijdelings van belang voor de selectie van een type zorginstelling om de stoornis van de verdachte te kunnen behandelen. Om te voorkomen dat een deskundige pro-justitia-rapporteur onterecht als tevens deskundig op het gebied van risicotaxatie wordt beschouwd en om te voorkomen dat een pro-justitia-rapporteur het stellen van een diagnose verwacht met het doen van een risicotaxatie, zou het dan wellicht aanbeveling kunnen verdienen om een onafhankelijke risicotaxateur de risicotaxatie te laten verrichten. Naar onze mening gaat dit echter te ver en kan de verantwoordelijkheid voor het verrichten van een risicotaxatie wel degelijk bij de pro-justitia-rapporteur worden neergelegd, maar zijn wellicht strengere eisen te stellen ten aanzien van het verrichten van risicotaxaties. Daarbij achten wij tevens een kritische houding van de rechter van belang.

Omdat de rechter geen diagnosticus is of deskundig op het gebied van individuele psychodiagnostiek, is het niet geëigend om de rechter een risicotaxatie te laten doen. Wel is het volgens ons mogelijk dat de rechter gebruikmaakt van informatie omtrent statische factoren – waarbij deze kan zijn verzameld vanuit een algoritmisch instrument en vanuit een GKO-instrument – en de veranderbare indicatoren tijdens de rechtszitting nader uitvraagt, zodat de rechter omtrent de invulling van deze items zowel informatie van de verdachte krijgt als van de pro-justitia-rapporteur. Om meer zekerheid over de persoonlijke omstandigheden in de klinische items te krijgen, zou het daarnaast aanbeveling verdienen om hierbij ook de reclassering, de verdachte, de naasten van de verdachte en (de naasten van) het slachtoffer in deze procedure te betrekken. Op deze wijze kan de rechter dan de informatie en argumentatie van veel naastbetrokkenen op waarde inschatten en wegen in het eindoordeel. Hierdoor heeft de rechter meer informatie, maar dit betekent helaas niet automatisch dat dan ook een betere

beoordeling wordt gemaakt, omdat voorspellen een lastige kwestie blijft.

Op naar een best practice

In het voorgaande hebben wij betoogd dat er veel vragen zijn omtrent het verrichten van risicotaxaties. Wij hebben besproken dat uit de recente meta-analyses blijkt dat het gemiddelde voorspellingsvermogen van algoritmische instrumenten en GKO-instrumenten niet veel van elkaar verschilt voor de voorspelling van een geweldsdelict of enig delict¹⁶ en dat er ook zeker geen overweldigend bewijs is dat dit anders is bij zedendelicten. Daarom zou men kunnen stellen dat het weinig uitmaakt welk type risicotaxatie wordt gebruikt. Echter ondanks het recente pleidooi van anderen voor het gebruik van een algoritmische risicotaxatie (vooral bij zedendelicten), hebben wij betoogd dat het GKO de preferente methode is vanwege de aansluiting op de rechtspraak en ook het behandelparadigma in de forensische psychiatrie. Hoewel ook de Stable-2007 richting biedt aan behandelingen, moet het ook bij zedendelicten ons inziens mogelijk zijn voor een onderzoeker of behandelaar om beredeneerd af te wijken van de uitkomsten van de algoritmische instrumenten en dus een gestructureerd klinisch oordeel te geven. Problematisch blijft echter dat het GKO net als de algoritmische methode, ook niet de mogelijkheid biedt om zeer betrouwbare en zeer nauwkeurige uitspraken te doen over specifieke risico's op korte, middellange of lange termijn en dat de uitspraken zich in het algemeen helaas nog moeten beperken tot het risico op algemene criminaliteit, een geweldsdelict of een zedendelict, zonder daarbij in te kunnen gaan op de precieze aard van de criminaliteit of het geweld of de omstandigheden waarin deze kunnen optreden. Nog problematischer is wellicht dat de inschattingen mogelijk weinig richting kunnen geven aan besluiten van de rechter, omdat elke beslissing betekent dat een groep mensen onjuist wordt bejegend. Bij hoog-risico-inschattingen is er bij 48 procent van de hoog scorende geweldplegers (HKT-R) geen sprake van recidive met een geweldsdelict en bij 60 procent van de hoogst scorende zedendelinquenten (Static-99R) geen nieuwe aanklacht of veroordeling voor een nieuw zedendelict. Oplegging van een stevige maatregel of detentie op basis van het ingeschatte risico zou dan bij deze mensen een te zware maatregel kunnen zijn. Andersom is er bij laag-risico-inschattingen bij 16 procent van de laag scorende geweldplegers en 1-3 procent van de laagst scorende zedendelinquenten wel degelijk sprake van recidive. Hoewel de absolute aantallen onjuiste beslissingen hierbij kunnen meevallen, hebben deze beslissingen op basis van het ingeschatte risico dus helaas een onterecht negatief effect voor de delictplegers of mensen in de maatschappij. Naar onze mening is rapportage over risicotaxaties te verbeteren door bij iedere risicotaxatie weer te geven welke risicofactoren zijn onderkend vanuit welk risicotaxatie-instrument, wat naast het ingeschatte risico (laag, midden, hoog) het daarbij behorende recidivecijfer is over bijvoorbeeld vijf jaar, en welke specifiek benoemde aanvullende inschattingen en overwegingen hebben geleid tot aanpassing van de oorspronkelijk algoritmische

inschatting. Deze wijze van communicatie bevordert inzicht in de achtergronden van inschattingen. Ook is naar onze mening de rapportage over risicotaxaties te verbeteren door duidelijk weer te geven op welke situaties de inschattingen betrekking hebben (bijvoorbeeld tijdens behandeling of toezicht) of waarom dit moeilijk kan worden beoordeeld.

We hebben betoogd dat de pro-justitia-rapporteur naar onze mening de geëigende persoon is voor het doen van de risicotaxaties, ook als geen stornis wordt vastgesteld omdat instrumenten meer gebaseerd zijn op karakterclusters en op symptoomclusters dan op stoornissen, maar dat het van belang is dat de rechter een kritische en onderzoekende houding aanneemt ten aanzien van de risicotaxatie. Het is duidelijk dat risicotaxatie-instrumenten nog veel onderzoek en verbetering behoeven, maar het is mogelijk om door multidisciplinair overleg of het bevragen van verschillende betrokkenen, zorgvuldiger inzicht te krijgen omtrent de waarschijnlijkheid van inschattingen en de juistheid van oordelen.

Referenties

- D.A. Andrews, J. Bonta & R.D. Hoge, 'Classification for effective rehabilitation: Rediscovering psychology', *Criminal Justice and Behavior* 1990, issue 17, p. 19-52.
- S. Bogaerts, M. Spreen, P. ter Horst & C. Gerlisma, 'Predictive validity of the HKT-R Risk Assessment for two and five-year recidivism in a cohort of Dutch Forensic Psychiatric Patients', *International Journal of Offender Therapy and Comparative Criminology* 2018, issue 62, p. 2259-2270.
- J. Bonta, M. Law & R.K. Hanson, 'The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis', *Psychological Bulletin* 1998, issue 123, p. 123-142.
- S. Fazel, J. Singh, H. Doll & M. Grann, 'Use of risk assessments instruments to predict violence and antisocial behaviour in 73 samples involving 24827 people: Systematic review and meta-analysis', *British Medical Journal* 2012, issue 345, p. e4692.
- W.M. Grove, D.H. Zald, B.S. Lebow, B.E. Snitz & C. Nelson, 'Clinical versus mechanical prediction: A meta-analysis', *Psychological Assessment* 2000, issue 12, p. 19-30.
- L.S. Guy, *Performance Indicators of the Structured Professional Judgment Approach for Assessing Risk for Violence to Others: A Meta-Analytic Survey* (diss. Burnaby), Burnaby, Canada: Simon Fraser University 2008.
- R.K. Hanson & M.T. Bussière, 'Predicting relapse: A meta-analysis of sexual offender studies', *Journal of Consulting and Clinical Psychology* 1998, issue 66, p. 348-362.
- R.K. Hanson & K. Morton-Bourgon, *Predictors of sexual recidivism: An updated meta-analysis* (Cat. No. PS3-1/2004-2E-PDF), Ottawa, ON: Public Works and Government Services Canada 2004.
- R.K. Hanson & K. Morton-Bourgon, 'The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies', *Psychological Assessment* 2009, issue 21, p. 1-21.
- J. Harte & M. Breukink, 'Objectiviteit of schijnzekerheid? Kwaliteit, mogelijkheden en beperkingen van instrumen-

16. Singh, Grann & Fazel 2011.

ten voor risicotaxatie', *Tijdschrift voor Criminologie* 2010, afl. 52, p. 52-72.

P.E. Meehl, *Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence*, Minneapolis: University of Minnesota Press 1954.

Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen, *Gewogen risico. Deel 1: Communiceren over recidive in zedenzaken*, Den Haag: Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen 2017.

C.E. van der Put & M. Assink, 'Misvattingen in de Nederlandse visie op risicotaxatie', *Tijdschrift voor Psychiatrie* 2018, afl. 60, p. 508-510.

J.P. Singh, M. Grann & S. Fazel, 'A comparative study of violence risk assessment tools: a systematic review and meta-regression analysis of 68 studies involving 25,980 participants', *Clinical Psychology Review* 2011, issue 31, p. 499-513.

M. Spreen, E. Brand, P. ter Horst & S. Bogaerts, *Handleiding en Methodologische Verantwoording HKT-R, Historisch, Klinische en Toekomstige – Revisie*, Groningen: FPC Dr. S. van Mesdag 2014.

D. de Vogel, D. Ruiters, V. Beek & G. Mead, 'Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders', *Law and Human Behavior* 2004, vol. 28, issue 3, p. 235-251.

D. de Vogel, M. de Vries Robbé, W. van Kalmthout & C. Place, 'Risicotaxatie van geweld bij vrouwen: Ontwikkeling van de "female additional manual" (fam)', *Tijdschrift voor Psychiatrie* 2012, afl. 54, p. 329-338.

S.G.C. van Wingerden, 'Risicotaxatie bij het straffen: a risky business', *Strafblad* 2018, afl. 16, p. 27-32.

S.G.C. van Wingerden, M. Moerings & J.A. van Wilsem, *Recidiverisico en straftoemeting*, Leiden: SDU 2011.