

Risicotaxatie

Citation for published version (APA):

de Ruiter, C., & Hildebrand, M. (2017). Risicotaxatie. In P. van Koppen, J. W. de Keijser, R. Horselenberg, & M. Jelicic (Eds.), *Routes van het Recht: Over de rechtspsychologie* (1 ed., pp. 983-999). Boom Juridisch.

Document status and date:

Published: 01/01/2017

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Risicotaxatie

Corine de Ruiter & Martin Hildebrand

Thomas is midden 50 en verblijft, na eerdere opnames in twee andere tbs-klinieken, op de longstay-afdeling van Forensisch Psychiatrisch Centrum Veldzicht.¹ Zijn advocaat vraagt in 2007 om een contra-expertise. Thomas is het niet eens met de risicotaxatie die over hem is opgesteld. Hij vindt zichzelf niet delictgevaarlijk en wil zijn longstay-plaatsing voor de rechter aanvechten. Volgens het dossier dat zijn advocaat beschikbaar stelt, is Thomas in 1993 veroordeeld tot vier jaar gevangenisstraf en tbs met dwangverpleging voor brandstichting in een schuur en in een woning. Er is niemand gewond geraakt. In de periode voorafgaand aan de brandstichtingen waren er toenemende spanningen in zijn huwelijk en mislukte zijn bedrijf. Thomas vertelt dat hij in een overspannen toestand, uit wraak en frustratie, tot zijn delicten kwam.

Hoe dient een deskundige te werk te gaan om te komen tot een wetenschappelijk onderbouwd oordeel over het recidiverisico van Thomas?

De term risicotaxatie is inmiddels ingeburgerd binnen het juridische domein in Nederland. Risico verwijst naar een dreiging van een ongewenste, schadelijke uitkomst; in de strafrechtelijke context is dat meestal het risico van nieuwe delicten.² Risico is een meervoudig begrip dat verwijst naar het type delictgedrag (bijvoorbeeld seksueel geweld of partnermishandeling), de waarschijnlijkheid dat het delictgedrag zal optreden, de frequentie of duur van het gedrag, de ernst van de gevolgen van dat gedrag, en de acuutheid of tijdstermijn ervan. Risico is per definitie veranderlijk en contextafhankelijk, omdat het ontstaat en toe- of afneemt in specifieke situaties. Risicotaxatie is het proces van het verzamelen van informatie ten behoeve van besluitvorming over de hoogte van het risico van delictgedrag.³

In dit hoofdstuk wordt eerst de geschiedenis van het wetenschappelijk onderzoek naar risicotaxatie van delictgedrag beschreven; het meeste onderzoek richt zich op risicotaxatie van verschillende vormen van geweld, zoals seksueel geweld, stalking en partnergeweld. Vooral in de forensische geestelijke gezondheidszorg, en in het bijzonder binnen het kader van de tbs-maatregel,⁴ is de invoering van gestructureerde risicotaxatie-instrumenten in het afgelopen decennium geformaliseerd in wet- en regelgeving.

1. Zie De Ruiter (2010) en De Ruiter (2015b) voor een gedetailleerde analyse van deze casus. Zie uitgebreider over de longstay: hoofdstuk 53 over de terbeschikkingstelling.
2. Bernstein (1996).
3. Segal & Coolidge (2000).
4. Zie hoofdstuk 53 over terbeschikkingstelling.

Ook de reclassering heeft risicotaxatie-instrumenten geïmplementeerd, om de advisering aan de rechter meer te standaardiseren. Daarvan zullen wij ook enkele voorbeelden geven. De laatste jaren is er ook kritiek geuit op de risicotaxatie-instrumenten. Ten eerste leggen zij een sterke nadruk op risicofactoren en hebben zij weinig aandacht voor (potentieel) beschermende factoren. Een tweede kritiekpunt is het feit dat toepassing van een risicotaxatie-instrument bij een individuele delinquent gemakkelijk tot een fout-positief oordeel leidt, waardoor ten onrechte geconcludeerd wordt dat er sprake is van een hoog risico. Ten slotte gaan we in dit hoofdstuk beknopt in op de communicatie over risico's. De casus Thomas loopt als een rode draad door dit hoofdstuk. Aan de hand van deze casus wordt uitgelegd hoe verschillende typen risicotaxatie-methoden in de praktijk werken.

Risicotaxatie in historisch perspectief

Ongestructureerd klinisch oordeel

Tot midden jaren negentig van de vorige eeuw waren taxaties van het risico van gewelddadig gedrag gebaseerd op de klinische blik van de deskundige, een soort intuïtie waardoor de deskundige, zonder gericht onderzoek uit te voeren, al weet 'wat er aan de hand is'. De klinische blik is globaal en ongestructureerd, steunt volledig op de expertise en ervaring van de expert, en heeft geen wetenschappelijk fundament. Deze klinische, ongestructureerde risicotaxatie was notoir onbetrouwbaar, omdat verschillende deskundigen tot heel verschillende oordelen kwamen over het geweldsrisico van een individu. Ook is dit soort taxatie inaccuraat. Gaandeweg werd uit steeds meer onderzoek duidelijk dat klinische beoordelingen een erg grote foutenmarge hebben.⁵ Dit werd voor de ongestructureerde risicotaxatie helder samengevat door Monahan, een Amerikaanse psycholoog, in zijn inmiddels klassieke monografie *The clinical prediction of violent behavior*.⁶ Hij deed daarin ook een oproep tot wetenschappelijke onderbouwing van de risicotaxatie van geweld. In de originele pro Justitia-rapportage over Thomas, die dateert uit midden jaren negentig, is van een wetenschappelijke onderbouwing van de risicotaxatie geen sprake. In die periode werd in Nederland niet of nauwelijks gebruikgemaakt van wetenschappelijk onderbouwde risicotaxatie-instrumenten.

In het pro Justitia-rapport dat het Pieter Baan Centrum op 9 maart 1993 over Thomas uitbracht, komt het woord risicotaxatie niet voor. Het rapport staat bol van het psychoanalytisch jargon ('verdrongen agressie', 'rivaliteit jegens mannen', 'plaatsvervangende identificatie', 'onbewuste drijfveren') en Thomas zou lijden aan een karakterneurotische stoornis,⁷ waarmee hij de vele agressieve en seksuele impulsen uit zijn bewustzijn weet te houden. Als het echter slechter met hem gaat, dan komen die impulsen door de karakterafweer heen, en raakt zijn persoonlijkheid gedissocieerd. Over het recidiverisico gaan alleen de volgende zinnen: 'Gevaar voor herhaling is zeker te vrezen, aangezien betrokkene's persoonlijkheidsstructuur en zijn sociale netwerk niet veranderd zijn. (...) Wan-

5. Grove & Meehl (1996).

6. Monahan (1981).

7. Overigens kwam deze stoornis ook in 1993 niet voor in de gangbare classificatiesystemen voor psychiatrische ziekten, zoals de *Diagnostic and Statistical Manual of Mental Disorders* en de *International Classification of Diseases*.

neer hij weer is vastgelopen in zijn minderwaardigheid zal hij alleen door middel van een analoge “acting-out” zijn narcistisch evenwicht kunnen herstellen. (...) Derhalve willen wij Uw College adviseren betrokkene de maatregel van terbeschikkingstelling met verpleging op te leggen’ (pro Justitia-rapport van 9 maart 1993, p. 53).

Het pro Justitia rapport over Thomas lijkt een typisch voorbeeld van de klinisch-psychoanalytische hermeneutiek die in die jaren in de mode was bij het Pieter Baan Centrum.⁸ Volgens de hermeneutiek moest het delict worden onderzocht en begrepen in relatie tot de levensloop van de verdachte, in het bijzonder zijn of haar vroege jeugdervaringen, die al dan niet waren verdrongen. Vandaar dat we begrippen als ‘verdrongen agressie’ en ‘rivaliteit jegens mannen’ in het PBC-rapport over Thomas aantreffen. Volgens deze psychoanalytische benadering is wetenschappelijke toetsbaarheid en transparantie van secundair belang,⁹ terwijl dat voor wetenschapsfilosofen als Karl Popper en Imre Lakatos juist reden was om de psychoanalyse te kwalificeren als pseudowetenschap.¹⁰

Brandstichters werden bijvoorbeeld door psychoanalytici gezien als zeer gevaarlijke, perverse individuen bij wie sprake was van gestoorde seksuele impulsen met een zeer hoog recidiverisico. Harris en Rice¹¹ bekritiseerden die psychoanalytische ideeën al in 1984, omdat zij louter op casuïstiek gebaseerd waren, terwijl empirisch onderzoek aantoonde dat brandstichters juist een lager recidiverisico hadden vergeleken met andere typen geweldsdelinquenten.

Actuariële risicotaxatie

In de jaren tachtig en negentig begonnen verschillende onderzoeksgroepen studie te doen naar de predictie van gewelddadig gedrag en andere vormen van criminaliteit, waaruit de actuariële risicotaxatie-instrumenten (ARAI's) zijn ontstaan. Bij actuariële risicotaxatie worden gegevens over risicofactoren verzameld, waarvan in empirisch onderzoek is gebleken dat die factoren significant samenhangen met het (opnieuw) plegen van gewelddadig gedrag. Volgens een vaststaande rekenformule (algoritme) worden de scores op de individuele factoren ‘omgezet’ in een conclusie over recidiverisico.¹² ARAI's beschouwen risico alleen in termen van de waarschijnlijkheid dat het delict optreedt, zonder rekening te houden met andere facetten van het risico, zoals aard, ernst en duur ervan. ARAI's werden ontwikkeld op basis van gegevens over groepen recidiverende en niet-recidiverende delinquenten, om een schatting te maken van de kans dat iemand in de toekomst zal recidiveren, gegeven bepaalde kenmerken. Dit type onderzoek was betrekkelijk eenvoudig van opzet: men keek in de dossiers van ex-gevangenen en codeerde een aantal historische factoren, zoals de leeftijd waarop het eerste misdrijf werd gepleegd, de sekse, en of er sprake was van een psychiatrische aandoening. Vervol-

8. Mooij (2006).

9. Mooij (2005).

10. Popper (1990) en Lakatos (1980).

11. Harris & Rice (1984).

12. De term ‘actuariële’ is afkomstig uit de verzekeringswereld. De actuaaris is daar degene die bij het afsluiten van een verzekering op grond van statistische gegevens onderzoekt wat de kansen zijn dat verzekerden uitkeringen zullen claimen. De hoogte van de premie wordt op dat gegeven gebaseerd, zodat er voldoende zekerheid is dat de verzekeringmaatschappij de claims kan dekken.

gens werd onderzocht welke van die historische factoren significant correleerden met het plegen van nieuwe delicten na ontslag uit de gevangenis. Op grond van hun voorspellende waarde werden gewichten aan de diverse factoren toegekend. Zo ontstonden instrumenten zoals de *Violence Risk Appraisal Guide* (VRAG)¹³ (zie tabel 1) en de *Ontario Domestic Assault Risk Assessment* (ODARA),¹⁴ die voor elke individuele delinquent berekenen tot welke risicocategorie (van laag tot hoog) ze behoren.

De meeste actuariële risicotaxatie-instrumenten bevatten alleen historische, statische risicofactoren, dat zijn kenmerken die in principe onveranderbaar zijn, zoals (jonge) leeftijd ten tijde van het indexdelict, een geschiedenis van problematisch alcoholgebruik of een geschiedenis van verblijf in tehuizen en pleeggezinnen. Dat komt doordat delictgedrag in de toekomst veruit het sterkst voorspeld wordt door kenmerken uit het verleden, in het bijzonder de criminele voorgeschiedenis.¹⁵ De kracht van actuariële instrumenten ligt in hun redelijk goede voorspellende waarde en hun relatief grote betrouwbaarheid (verschillende beoordelaars komen met het betreffende instrument tot vergelijkbare oordelen over dezelfde persoon). Een belangrijke beperking van de actuariële benadering is de afwezigheid van dynamische, veranderbare risicofactoren in de instrumenten. Dynamische risicofactoren zijn kenmerken die in principe kunnen veranderen, bijvoorbeeld door behandeling of louter door het verstrijken van de tijd, zoals impulsiviteit en zelfinzicht.

De pro Justitia-rapporteurs die over Thomas advies uitbrachten aan de rechter, vonden op basis van hun klinisch oordeel dat hij een hoog risico vormde op het opnieuw plegen van brandstichting en zij adviseerden de maatregel van tbs met dwangverpleging. Wat was de uitkomst van de risicotaxatie geweest als de deskundigen een actuariel risicotaxatie-instrument hadden kunnen gebruiken en hun oordeel daarop hadden gebaseerd?

In 1993 was er geen specifiek actuariel risicotaxatie-instrument voor brandstichters. Studies laten zien dat brandstichters risicofactoren gemeen hebben met gewelddadige en met niet-gewelddadige delinquenten.¹⁶ Quinsey en collega's vonden dat door gebruik te maken van dezelfde statistische methode waarmee de VRAG eerder was ontwikkeld, uiteindelijk zeven risicofactoren tot een optimale voorspelling van brandstichting in de toekomst leiden: jonge leeftijd bij eerste brandstichting, totale aantal brandstichtingen, brandstichting als kind, laag IQ, andere delicten samengaand met de brandstichting, de brand alleen stichten, en een lage agressiescore.¹⁷ Quinsey en collega's rapporteren dat de totaalscore van de lijst van die zeven factoren een significante voorspelling van brandstichting opleverde, met een ROC-AUC-waarde van ,76.¹⁸

13. Quinsey, Harris, Rice & Cormier (1998). Nederlandse vertaling: Jeandarme, Pouls & Peters (2012).
14. Hilton, Harris, Rice, Lang & Cormier (2004).
15. Klassen & O'Connor (1994) en Monahan (1981).
16. Edwards & Grace (2014) en Rice & Harris (1996).
17. Een lage agressiescore lijkt tegen-intuïtief, maar het blijkt dat brandstichters voornamelijk behoren tot de categorie 'oppotters' van boosheid. Het zijn vaak sub-assertieve personen die moeite hebben om (normale) boosheid in interpersoonlijke relaties te uiten.
18. ROC-AUC staat voor *Area Under the Curve* van de *Receiver Operating Characteristics analyse*. De diagnostische nauwkeurigheid van risicotaxatie-instrumenten wordt meestal uitgedrukt in termen van de AUC-waarde. Deze statistische grootheid is afkomstig uit de signaal-detectietheorie (denk aan het testen van de precisie van radarsystemen). Kort gezegd geeft de ROC weer hoe de verhouding is tussen fout-positieve en fout-negatieve voorspellingen bij elke

Toch ontwikkelden Quinsey en collega's uiteindelijk geen separaat risicotaxatie-instrument voor brandstichters, maar raden zij aan de VRAG te gebruiken (zie tabel 1 voor de scoring van de VRAG voor Thomas). Thomas behaalt een totaalscore van 0 op de VRAG. Volgens de VRAG-handleiding hoort bij die score de risicocategorie Matig, wat overeenkomt met een recidivekans van 35% over een periode van zeven jaar na vrijlating.¹⁹

Tabel 1: VRAG Thomas

Risicofactor	Score
Geen scheiding meer dan 1 maand van minstens 1 ouder of geen traumatische scheiding	-2
Geen aanwezigheidsproblemen en/of probleemgedrag op school	-1
Geen geschiedenis van problematisch alcoholgebruik	-1
Burgerlijke staat tot en met indexdelict = ooit gehuwd	-2
Delictgeschiedenis niet-gewelddadige delicten	+3
Schending voorwaarden tot en met het indexdelict	+3
Leeftijd ten tijde van indexdelict	-1
Geen letsel slachtoffer indexdelict	+2
Geen slachtoffer	+1
Geen persoonlijkheidsstoornis	-2
Geen schizofrenie	+1
PCL-R-score (10-14)	-1
Totaalscore VRAG	0

Base rate

Een belangrijk aspect bij risicotaxatie is kennis over de *base rate* van recidive, dat wil zeggen kennis over de statistische prevalentie (frequentie) van recidive in een gegeven groep (bijvoorbeeld tbs-gestelden), over een bepaalde tijdperiode (bijvoorbeeld 1 jaar of 5 jaar).²⁰ De *base rate* is afhankelijk van meerdere factoren, zoals de lengte van de observatieperiode, het recidivecriterium (bijvoorbeeld arrestatie, veroordeling, detentie, zelfrapportage) en de aard van de recidive (bijvoorbeeld: niet-seksueel, met slachtoffer; niet-seksueel, geen slachtoffer; seksueel, met slachtoffer). De hoogte van de *base rate* heeft invloed op de mogelijke verbetering van de risicotaxatie ten opzichte van **kan-****sniveau** door gebruik van een risicotaxatie-instrument. Wanneer de *base rate* erg laag (bijvoorbeeld 5%) of erg hoog is (bijvoorbeeld 95%), heeft het uitvoeren van een gestructu-



mogelijke afkapwaarde van het instrument. Door voor elk afkappunt de percentages 'raak' en 'fout' tegen elkaar af te zetten, ontstaat een curve. Bij risicotaxatie betreft het correct voorspelde recidives tegenover recidives die zich uiteindelijk niet hebben voorgedaan (de 'valse alarmen'). De oppervlakte onder die curve, de AUC, is een maat voor de algemene voorspingskwaliteit van het instrument. Een waarde van 0,50 is gelijk aan een voorspelling die niet beter is dan toeval, een waarde van 0,70 wordt doorgaans gezien als 'redelijk goed' en 1,00 betekent een perfecte voorspelling. Het verschil tussen 0,50 en de feitelijk gevonden waarde (bijvoorbeeld 0,70) geeft aan hoeveel winst er geboekt wordt ten opzichte van toeval (kans) door een risicotaxatie-instrument te gebruiken.

19. Jeandarme, Pouls & Peters (2012), p. 32.

20. Cunningham & Reidy (1998).

reerde risicotaxatie (dat wil zeggen: met behulp van een instrument) in het algemeen weinig zin. Wanneer bijvoorbeeld maar 5% in een bepaalde populatie recidiveert, doen we al een uitstekende ‘voorspelling’ als we van iedereen in die betreffende populatie zeggen – zonder gebruik te maken van wat voor instrument dan ook – dat hij *niet* zal recidiveren. In het geval van een erg hoge *base rate* (bijvoorbeeld 95%) ‘voorspellen’ we uitstekend als we van iedereen in de betreffende populatie zeggen – wederom zonder instrument – dat hij *wél* zal recidiveren. Voor beide voorbeelden geldt dat een risicotaxatie-instrument extreem accuraat en betrouwbaar moet zijn om nog iets aan de voorspelling op basis van kans te kunnen verbeteren. Wanneer de *base rate* dicht(er) bij 50% ligt, heeft gebruik van een risicotaxatie-instrument wel zin, omdat de toevalskans in die gevallen geen houvast biedt; die levert in de helft van de gevallen goede dan wel foute voorspellingen op (vergelijkbaar met het opgooien van een munt). In dergelijke gevallen zal het gebruik van een risicotaxatie-instrument relatief eenvoudig tot verbetering van accuraatheid van de taxatie kunnen leiden.²¹ In de (risicotaxatie)praktijk wordt het belang van de *base rate* vaak onvoldoende onderkend. Dat verschijnsel wordt wel aangeduid als de *base rate fallacy*. Kennis van de *base rate*-statistieken van de groep waarbinnen een risicotaxatie moet worden gedaan is van belang, omdat clinici vaak weinig oog hebben voor de *base rates* van specifiek gewelddadig gedrag en geneigd zijn tot overpredictie bij het schatten van delict risico's. Bij de actuariële risicotaxatiebenadering is de *base rate* van recidive verdisconteerd in de indeling in de verschillende risicocategorieën (bijvoorbeeld ‘laag’, ‘matig’ en ‘hoog’ risico). We illustreren dit aan de hand van tabel 2, waarin recidivecijfers van 242 mannelijke tbs-gestelden staan vermeld, voor een periode van 5,2 jaar na afloop van de tbs-maatregel. De gegevens zijn afkomstig uit de handleiding van de HKT-R, een Nederlandstalig risicotaxatie-instrument voor geweldsdelicten.

Tabel 2: Berekening van de diagnostische waarde van de HKT-R op basis van onderzoek onder 242 tbs-gestelden

		In werkelijkheid			
		Wel recidive		Geen recidive	
HKT-R	Hoog	22	38%	20	11%
	Matig	10	17%	28	15%
	Laag	26	45%	136	74%
		58	100%	184	100%

Noot. Cijfers overgenomen uit Spreen, Brand, Ter Horst & Bogaerts (2014).

De diagnostische waarde van een categorisering als Hoog risico volgens de HKT-R is het percentage correct geclassificeerd als hoog risico gedeeld door het percentage onterecht als hoog risico geclassificeerd (de fout-positieven).²² Dat levert hier op: $38/11 = 3,45$. De diagnostische waarde voor werkelijke recidive van de HKT-R-uitslag Hoog risico is dus 3,45. De diagnostische waarde van een classificatie Laag risico voor geen recidive is $74/45 = 1,64$.

21. Vergelijk bijvoorbeeld Grubin & Wingate (1996).

22. Zie voor meer uitleg over het begrip diagnostische waarde hoofdstuk 20 over de waarde van bewijs.

Hoewel de voorspellende waarde van actuariële risicotaxatie-instrumenten in de meeste studies significant beter is dan kansniveau, en ook significant beter dan die van het ongestructureerde klinisch oordeel,²³ is ook de actuariële benadering niet zonder kritiek gebleven.²⁴ Ten eerste houdt de actuariële benadering geen rekening met het feit dat er naast de statische, historische risicofactoren, ook dynamische, veranderbare factoren zijn die medebepalend kunnen zijn voor het al dan niet optreden van recidive. De meeste actuariële risicotaxatie-instrumenten bestaan uit een verzameling van tien tot twaalf risicofactoren die statisch zijn en dus ook niet ten goede kunnen veranderen.²⁵ Ze kunnen een stigmatiserend effect hebben omdat een eenmaal vastgesteld hoog recidiverisico niet omlaag kan worden gebracht. Een tweede bezwaar is meer technisch van aard en houdt in dat de actuariële instrumenten op analyses voor groepen delinquenten gebaseerd zijn, maar dat die groepsgegevens niet zomaar toegepast mogen worden op het individuele geval. Uit diverse studies blijkt namelijk dat wanneer je een actuariële instrument toepast op een individu, de foutmarge van de risicoschatting zo groot is, dat er sprake is van schijnnaauwkeurigheid.²⁶

Gestructureerd professioneel oordeel

Bovenstaande kritiek op de actuariële benadering van risicotaxatie leidde in de jaren negentig tot een alternatieve benadering, die het gestructureerde professionele oordeel (*structured professional judgment; SPJ*) wordt genoemd. De risicotaxatie-instrumenten die tot deze benadering gerekend worden, onderscheiden zich op drie belangrijke punten van de actuariële instrumenten: (1) ze bevatten naast historische ook dynamische risicofactoren, (2) er is geen vaste wegging van de risicofactoren, en (3) er is ruimte voor toevoeging van casus-specifieke risicofactoren, en (4) er wordt rekening gehouden met de context waarvoor de risicotaxatie geldt. Inmiddels zijn er, net als bij de actuariële benadering, voor verschillende typen delicten specifieke SPJ-instrumenten beschikbaar, zoals de HCR-20^{v3} en de HKT-R voor geweld in het algemeen en de B-SAFER specifiek voor risicotaxatie bij relationeel geweld.²⁷ Naar de meeste SPJ-instrumenten is in Nederland onderzoek verricht, met over het algemeen vergelijkbare voorspellende waarden met die gevonden worden in buitenlands onderzoek.²⁸ We zullen de risicotaxatie volgens het SPJ model hier illustreren aan de hand van de casus Thomas.

Omdat er (nog) geen specifiek SPJ-instrument voor brandstichters bestaat, is gekozen voor afname van de HCR-20, een SPJ-instrument voor alle vormen van geweld. In gevallen waarbij de brandstichting gemotiveerd is vanuit boosheid en wraak, zoals bij

23. Ægisdóttir, White, Spengler, Maugherman, Anderson, Cook, Nichols, Lampropoulos, Walker, Cohen & Rush (2006) en Grove, Zald, Lebow, Snitz & Nelson (2000). Zie echter een recente meta-analyse van Fazel, Singh, Doll & Grann (2012), waarin geen significant verschil werd gevonden tussen de predictieve validiteit van historisch-actuariële instrumenten en ongestructureerde klinische oordelen.
24. Hart & Cooke (2013).
25. Rogers (2000) en Petrila & De Ruiter (2011).
26. Hart & Cooke (2013) en Henderson & Keiding (2005).
27. Douglas, Hart, Webster & Belfrage (2013) en Kropp, Hart & Belfrage (2005). Nederlandse vertalingen: De Vogel, De Vries Robbé, Bouman, Chakhssi & De Ruiter (2013c) en De Ruiter (2011).
28. Lodewijks, Doreleijers & De Ruiter (2008), De Vogel, De Ruiter, Hildebrand, Bos & Van de Ven (2004) en De Vogel, De Ruiter, Van Beek & Mead (2004).

Thomas, mag de HCR-20 gebruikt worden.²⁹ Geweld wordt in de HCR-20 gedefinieerd als 'daadwerkelijk toegebracht lichamelijke letsel aan een andere persoon, een poging daartoe of dreiging daarmee'.³⁰ De HCR-20 bestaat uit tien historische risicofactoren, vijf dynamische, klinische factoren en vijf risicomangementfactoren, die op de toekomst gericht zijn (zie tabel 3 voor de items van de HCR-20). Voor alle risicofactoren geldt dat zij empirisch onderzocht zijn en samenhangen met een verhoogd risico van geweld. Alle risicofactoren worden gecodeerd op een driepuntsschaal: 0 = de risicofactor is afwezig; 1 = de risicofactor is in enige/mindere mate aanwezig; 2 = de risicofactor is duidelijk aanwezig. Codering van de items van de HCR-20 bij Thomas levert een score 2 op voor de historische factoren H4: Problemen in het arbeidsverleden en H8: Problemen in de kindertijd. Een aantal historische risicofactoren is in mindere mate aanwezig en krijgt een score 1: H1: Eerder geweld, H3: Instabiliteit in relaties, H10: Eerdere onttrekking aan toezicht. De overige vier historische risicofactoren (H5: Problemen met middelengebruik, H6: Psychische stoornis, H7: Psychopathie, en H9: Persoonlijkheidsstoornis) zijn niet aanwezig. Van de vijf klinische risicofactoren die betrekking hebben op het functioneren in de afgelopen zes maanden, zijn er drie met een score 1: C1: Gebrek aan zelfinzicht, C4: Impulsiviteit en C5: Reageert niet op behandeling. Er is geen sprake van C3: Actieve psychotische symptomen of van C2: Negatieve opvattingen die geweld goedkeuren. De risicohanteringsfactoren worden gescoord met het oog op een toekomstig verblijf in de maatschappij. Drie van de vijf risicofactoren zullen bij terugkeer in de maatschappij van Thomas niet aanwezig zijn: R2: Blootstelling aan destabiliserende factoren, R3: Geringe beschikbaarheid van persoonlijke steun en R5: Hoog niveau van ervaren stress. De destabiliserende factoren die in de aanloop naar de oorspronkelijke brandstichtingen aanwezig waren, zoals stress door een eigen bedrijf, relatieproblemen, zijn niet meer aanwezig. Wat betreft sociale steun zijn de voorwaarden gunstig: Thomas heeft een goede relatie met zijn zus, die in de loop van zijn verblijf in de tbs versterkt is. Ook met zijn zoon heeft hij regelmatig en goed contact. De risicofactoren R1 en R4 kunnen niet worden gecodeerd, omdat daartoe eerst een concreet resocialisatieplan dient te worden gemaakt. Gezien zijn huidige verblijf op de longstay-afdeling ontbreekt zo'n plan. Dat betekent dat daarmee in het eindoordeel rekening dient te worden gehouden.

Wanneer we de codering van de HCR-20 overzien, kunnen we constateren dat de optelling van de scores op de aanwezige risicofactoren leidt tot een relatief lage totaalscore, namelijk 10 (bij een maximale score van 40). Het gestructureerde professionele eindoordeel op de HCR-20 is echter meer dan een simpele optelsom van de risicofactoren. We dienen ook te onderzoeken of er mogelijk sprake is van een combinatie of interactie van risicofactoren die risico-verhogend zou kunnen werken. Bij Thomas is dat niet het geval. Uit de onderzoeksliteratuur weten we dat recidive bij brandstichters van het wraakzuchtige type vooral samenhangt met alcohol- en drugmisbruik en met een antisociale persoonlijkheidsstoornis.³¹ Beide diagnoses zijn bij Thomas afwezig. Dat betekent dat op basis van de risicotaxatie uitgevoerd met de HCR-20 kan worden geconcludeerd dat er sprake is van een laag risico van recidive met brandstichting. Een laag

29. Webster, Douglas, Eaves & Hart (1997).

30. De Vogel, De Vries Robbé, Bouman, Chakhssi & De Ruiters (2013a), p. 17. De definitie van geweld in de HCR-20 is consistent met de definitie die door vele anderen wordt gehanteerd, bijvoorbeeld de definitie van interpersoonlijk geweld van de World Health Organization (Krug, Dahlberg, Mercy, Zwi & Lozano, 2002).

31. Lindberg, Holi, Tani & Virkkunen (2005) en Repo, Virkkunen, Rawlings & Linnoila (1997).

recidiverisico impliceert dat volstaan kan worden met beperkte vormen van toezicht en ondersteuning bij terugkeer in de maatschappij.

Tabel 3: Risicofactoren van de HCR-20

<i>Historisch (H-schaal)</i>	<i>Klinisch (C-schaal)</i>	<i>Risicohantering (R-schaal)</i>
H1 Eerder gewelddadig gedrag	C1 Gebrek aan zelfinzicht	R1 Geringe kans dat plannen zullen slagen
H2 Jonge leeftijd bij eerste geweldsincident	C2 Negatieve opvattingen	R2 Blootstelling aan destabiliserende factoren
H3 Instabiliteit van relaties	C3 Actieve psychotische symptomen	R3 Geringe beschikbaarheid sociale steun
H4 Problemen in het arbeidsverleden	C4 Impulsiviteit	R4 Werkt niet mee aan behandelmogelijkheden
H5 Problemen met middelengebruik	C5 Reageert niet op behandeling	R5 Hoog niveau van ervaren stress
H6 Psychische stoornis		
H7 Psychopathie		
H8 Problemen in de kindertijd		
H9 Persoonlijkheidsstoornis		
H10 Eerdere onttrekking aan toezicht		

Gebruik van risicotaxatie-instrumenten buiten de tbs-sector

Ook in aanpalende sectoren, zoals de reclassering en de ambulante forensische psychiatrie, zijn gestructureerde risicotaxatie-methoden ingevoerd. De Recidive Inschattingsschalen (RISc) is het diagnose-instrument voor de reclassering en het gevangeniswezen, waarmee reclasseringswerkers op gestructureerde wijze het algemene recidiverisico van delinquenten taxeren, met daaraan gekoppeld een advies over eventuele noodzakelijke interventies gericht op het verminderen van de risico's.³² In de praktijk wordt de RISc sinds november 2004 voor verschillende doelen gebruikt, bijvoorbeeld voor reclasseringsadviezen aan de rechter of het openbaar ministerie of voor het ontwikkelen van een plan van aanpak voor reclasseringstoezicht. De RISc is gebaseerd op het Britse *Offender Assessment System* (OASys; Home Office, 2002), een integraal systeem van diagnostiek, indicatiestelling en sanctieplanning, ontwikkeld door de *National Probation Services* (reclassering) en *National Prison Services* (gevangeniswezen) van Engeland en Wales.³³ De RISc richt zich, daarbij gesteund door de zogenaamde *What Works*-benadering van

32. Adviesbureau Van Montfoort & Reclassering Nederland (2004).

33. OASys is op zijn beurt ontwikkeld op basis van de *Level of Service Inventory-Revised* (LSI-R; Andrews & Bonta, 1996) en het *Assessment Case management and Evaluation System* (ACE; Gibbs, 1999). De LSI-R, later herzien als de *Level of Service/Case Management Inventory* (LS/CM; Andrews, Bonta & Wormith, 2004), wordt in Nederland nauwelijks gebruikt, maar behoort internationaal tot de best gevalideerde risicotaxatie-instrumenten.

crimineel gedrag,³⁴ zowel op het in kaart brengen van historische als dynamische risicofactoren, bij zowel de dader als zijn sociale omgeving, als op zijn mogelijkheden om daar verandering in aan te brengen (door interventies). In de huidige versie van de RISC (RISC 4), die sinds juni 2014 in gebruik is,³⁵ worden de volgende twaalf schalen onderscheiden:

1. Analyse aspecten huidige delict
2. Delictgeschiedenis
3. Huisvesting en wonen
4. Opleiding, werk en leren
5. Inkomen en omgaan met geld
6. Relaties met partner, gezin en familie
7. Relaties met vrienden en kennissen
8. Drugsgebruik
9. Alcoholgebruik
10. Geschiedenis van antisociale gedragingen
11. Denkpatronen, gedrag en vaardigheden
12. Procriminele houding

De eerste twee schalen hebben betrekking op het huidige delict of de huidige tenlastelegging (schaal 1) en de delictgeschiedenis van de delinquent (schaal 2). Op deze schalen, alsmede op schaal 10, worden de meeste historische risicofactoren in kaart gebracht. Op de andere schalen wordt informatie verzameld over dynamische, in principe veranderbare, risicofactoren. Elke schaal bestaat uit meerdere items. De reclasseringswerker dient aan de hand van de scoringsinstructies bij de items een score te geven.³⁶ Zo ontstaat op elke schaal, met uitzondering van schaal 1,³⁷ een schaalscore. De schaalcores geven een indicatie van de mate waarin de risicofactor bijdraagt aan het delictgedrag van de delinquent; hoe hoger de score, hoe groter de invloed van de factor op het delictgedrag. De RISC genereert, op basis van een deel van de scores (de schalen 2 tot en met 7), een op empirische gegevens gebaseerde automatische inschatting³⁸ van het recidi-

34. Zie voor een bespreking van de *What Works*-principes hoofdstuk 10.

35. Hildebrand & Reclassering Nederland (2014). Het onderzoek naar de psychometrische kwaliteiten van de eerste versie van de RISC en ervaringen van reclasseringswerkers heeft geleid tot verdere ontwikkeling van RISC. Dat heeft uiteindelijk geresulteerd in RISC-versie 4 (RISC 4).

36. De te scoren items zijn zo veel mogelijk gebaseerd op concrete kenmerken van de delinquent of zijn omstandigheden, en hebben betrekking op onderwerpen waarvan uit wetenschappelijk onderzoek is gebleken dat deze verband houden met de kans op recidive.

37. In schaal 1 beschrijft en analyseert de reclasseringswerker de achtergronden en motieven van het huidige delict. Bij de analyse van het delict (de delicten) wordt gekeken naar bijkomende factoren en omstandigheden (zoals gebruik van wapens, mate van geweld, slachtoffers/nabestaanden, seksuele aspecten, et cetera), en naar de aanleiding tot het delict. Dat is moeilijk in scores te vatten.

38. De automatische inschatting van het recidiverisico in recidivecategorieën in de RISC 4 is gebaseerd op de totaalscore op een verzameling van 27 items van de schalen 2 tot en met 7 (somscore 0-54). Dat zijn de items die zowel in de eerste versie van de RISC als in RISC 4 zijn opgenomen, zonder al te grote inhoudelijke aanpassingen. Aan de hand van een dataset van het WODC (16.211 reclasseringscliënten) is de relatie tussen de score op bovengenoemde set van 27 items en recidive binnen twee jaar berekend en op basis van de relatie somscore en recidive een indeling in recidivecategorieën gemaakt. Dit is een tijdelijke oplossing. Zodra

verisico in vier risicocategorieën: 'laag risico', 'matig risico', 'hoog risico' en 'zeer hoog risico'. Het is vervolgens de taak van de reclasseringswerker om kritisch te beoordelen of het noodzakelijk is om op basis van het *gehele* profiel, dus inclusief de schalen 8 tot en met 12, die nog geen deel uitmaken van de automatisch gegenereerde risico-inschatting, af te wijken van de automatische risico-inschatting. Het eindoordeel over het recidiverisico is in die zin dus (ook) een gestructureerd professioneel oordeel, waardoor de RISc een zogenaamd hybride instrument wordt (deels actuarieel, vanwege de automatische recidive-inschatting, deels gestructureerd professioneel oordeel).

Naast het verzamelen van informatie over criminogene risicofactoren wordt in de RISc ook ingegaan op de responsiviteit van de cliënt.³⁹ De responsiviteitsfactoren krijgen geen score, maar de informatie is wel belangrijk bij de keuze voor een interventie (het gaat bijvoorbeeld om iemands motivatie, gezondheidsproblemen, of zijn taalachterstand). De mening van de delinquent zelf komt aan bod in de zelfrapportage van de RISc, een korte vragenlijst waarin de leefgebieden die in de RISc gescoord worden, terugkomen en waarin wordt nagevraagd hoe de delinquent zelf tegen zijn problematiek aankijkt. Ook wordt er in de RISc aandacht gevraagd voor beschermende of protectieve factoren, kenmerken van de delinquent of zijn omstandigheden die het delictrisico verkleinen.

Inmiddels is er enig onderzoek naar de psychometrische eigenschappen en voorspellingskracht van de RISc beschikbaar.⁴⁰ Onderzoek naar de interbeoordelaarsbetrouwbaarheid, schaalstructuur en interne consistentie van de RISc heeft aangetoond dat die in het algemeen voldoende zijn,⁴¹ maar dat op onderdelen verbeteringen nodig zijn.⁴² De mate waarin de RISc een goede inschatting geeft van het recidiverisico, is eveneens onderzocht. De totaalscore van de RISc laat een significante, redelijk goede voorspellende waarde zien voor algemene recidive, ROC-AUC = .70 ($N = 13.364$).⁴³ Voor het voorspellen van specifieke vormen van recidive, zoals gewelddadige recidive of seksuele recidive, was de voorspellende waarde van de RISc onvoldoende.⁴⁴ De RISc voorspelt niet alleen algemene recidive redelijk goed, het instrument blijkt ook redelijk in staat te voorspellen welke mannelijke reclasseringscliënten zich onttrekken aan de reclasseringsvoorwaarden, ROC-AUC = .70 ($N = 14.363$).⁴⁵

Op basis van tabel B4c uit het rapport van Van der Knaap en Alberda kan de diagnostische waarde van de RISc worden berekend voor algemene, gewelddadige en seksuele recidive (binnen twee jaar): Algemene recidive: Hoog risico = 2,78 (26,4/9,5) en Laag risico = 2,24 (50,3/22,5); gewelddadige recidive: Hoog risico = 1,87 (26,2/14,0) en

er voldoende gegevens zijn voor een koppeling van alle schalen aan recidivegegevens, zal de automatische inschatting van het recidiverisico worden bepaald op basis van alle schalen.

39. Responsiviteit gaat over de persoon van de dader, zijn intellectuele en sociale capaciteiten, zijn leerstijl en zijn motivatie. Maar het gaat ook over de behandelaar, trainer, toezichthouder en de soort interventie. Het 'responsiviteitsbeginsel' betekent dat het (behandel)aanbod moet passen bij de individuele delinquent.
40. Het betreft onderzoek dat met de eerste versie van de RISc (2004) is uitgevoerd.
41. Van der Knaap, Leenarts & Nijssen (2007).
42. Meer daarover in Hildebrand & Bosker (2011).
43. Van der Knaap & Alberda (2009).
44. Dat is op zich niet verbazingwekkend, de RISc is daar ook niet voor ontworpen.
45. Hildebrand, Hol & Bosker (2013).

Laag risico = 2,07 (43,1/20,8); seksuele recidive: Hoog risico = 1,18 (18,2/15,4) en Laag risico = 1,41 (40,6/28.8).⁴⁶

Problemen met actuariële en SPJ-instrumenten

Uit de casus Thomas wordt duidelijk dat het oorspronkelijke klinische oordeel 'hoog risico' van de pro Justitia-rapporteurs uit midden jaren negentig niet overeenstemt met de uitslag van beide 'modernere' benaderingen van risicotaxatie. Een objectievere, op empirische leest geschoeide, risicotaxatie leidt in dit geval tot een ander oordeel dan het subjectieve oordeel van de deskundigen. Toch kleven ook aan de moderne risicotaxatie-methoden nadelen.

Eerder in dit hoofdstuk werd al gewezen op de bezwaren die door sommige auteurs opgeworpen zijn tegen het gebruik van actuariële risicotaxatie-instrumenten voor de predictie van het recidiverisico in individuele gevallen, omdat hun diagnostische nauwkeurigheid beperkt is.⁴⁷ De diagnostische nauwkeurigheid van risicotaxatie-instrumenten wordt meestal uitgedrukt in termen van de zogenaamde AUC-waarde. Dat is de kans dat een willekeurig gekozen delinquent die is gerecidiveerd, een hogere score heeft op het instrument dan een *at random* gekozen delinquent die niet recidiveerde. De AUC is in de afgelopen decennia *de facto* de standaard geworden in het onderzoek naar de voorspellende waarde van risicotaxatie-instrumenten, maar de AUC is een nogal grove en generieke maat, die voor de forensische praktijk van weinig waarde is.⁴⁸ Een ander bezwaar tegen de AUC is dat het een schatting van de predictieve validiteit oplevert die onafhankelijk is van de *base rate*, terwijl in het forensische veld de klinische bruikbaarheid van een instrument juist afhankelijk is van de *base rate*. Zie het voorbeeld van de HKT dat wij eerder gaven in dit hoofdstuk (tabel 2).

Twee maten voor predictieve nauwkeurigheid die relevanter zijn voor de forensische risicotaxatie, zijn de Positieve Predictieve Waarde (*Positive Predictive Value*; PPV) en de Negatieve Predictieve Waarde (*Negative Predictive Value*; NPV). De PPV is de proportie van degenen die als hoog risico beoordeeld zijn op basis van het risicotaxatie-instrument, die daarna daadwerkelijk een nieuw delict pleegt, terwijl de NPV verwijst naar de proportie van hen die als laag risico zijn beoordeeld, die geen nieuw delict pleegt.⁴⁹ Een instrument met een hoge PPV kan dan worden geselecteerd om in de forensische praktijk hoog-risicogevoallen te detecteren; een instrument met een hoge NPV kan worden gebruikt om laag-risicogevoallen uit te filteren. Uit een recente meta-analyse die 73 studies omvatte met in totaal bijna 25.000 delinquenten, bleek echter dat de meeste risicotaxatie-instrumenten weliswaar een hoge NPV hebben, maar dat de hoogte van de PPV vaak te wensen overlaat (zie tabel 4).⁵⁰ Met andere woorden, veel personen die als matig of hoog risico worden getaxeerd met risicotaxatie-instrumenten, hebben in feite een laag risico op recidive. Hildebrand en collega's vonden bijvoorbeeld dat van de tbs-gestelden die op basis van de HCR-20 waren beoordeeld als hoog risico, 62% niet recidiveerde met een gewelddadig delict.⁵¹ Ook De Vogel vond dat een aanzienlijk percentage (tussen 38

46. Van der Knaap & Alberda (2009), p. 92.

47. Hart & Cooke (2013).

48. Singh (2013).

49. Altman & Bland (1994).

50. Fazel, Singh, Doll & Grann (2012).

51. Hildebrand, Hesper, Spreen & Nijman (2005).

en 50%) van de tbs-patiënten die op basis van de HCR-20 werden beoordeeld als hoog risico, niet opnieuw een gewelddadig delict pleegde.⁵² Met andere woorden, de instrumenten die richtinggevend zijn in het tbs-veld, overschatten risico's: laag-risico-individen worden adequaat geïdentificeerd, maar het aanwijzen van de hoog-risicogevallen gebeurt veel minder goed. Deze resultaten maken duidelijk dat (ook) moderne risicotaxatie-instrumenten in de praktijk nog steeds relatief veel vals-positieven opleveren.

Tabel 4: Samenvatting van enkele validiteitsmaten voor risicotaxatie-instrumenten voor drie typen delicten (geweld, seksueel geweld en algemene criminaliteit)

Validiteitsmaat	Geweld (n = 30) [*]	Seksueel geweld ^{**}	Criminaliteit ^{***}
AUC	.72	.74	.66
PPV	.41	.23	.52
NPV	.91	.93	.76

^{*} HCR-20, SARA, SAVRY en VRAG.

^{**} SORAG, Static-99 en SVR-20.

^{***} LSI-R en PCL-R.

Overgenomen uit: Fazel, Singh, Doll & Grann (2012)



Beschermende factoren

De implementatie van gestructureerde instrumenten voor risicotaxatie binnen de forensische sector in Nederland werd zeker niet door alle professionals vanaf het begin af aan omarmd. Er was aarzeling en soms zelfs regelrechte weerstand, omdat men vond dat daarmee de nadruk te veel kwam te liggen op de negatieve kenmerken van de delinquenten. Die mening werd ook helder verwoord door Rogers in een kritisch artikel uit 2000 over de stand van het risicotaxatie-onderzoek van dat moment.⁵³ Rogers vond dat de extreme nadruk op risicofactoren het gevaar in zich hield van pessimisme onder therapeuten, stigmatisering van delinquenten, en uiteindelijk tot onnodig lange detentie van forensisch psychiatrische patiënten zou leiden. Rond diezelfde tijd startten wij in de Van der Hoeven Kliniek met de ontwikkeling van een instrument dat beschermende factoren meet, als uitbreiding op de bestaande risicotaxatie-instrumenten, zoals de HCR-20 en SVR-20.⁵⁴ Dit instrument, de SAPROF (*Structured Assessment of PROtective Factors for violence risk*),⁵⁵ bevat zeventien beschermende factoren, ingedeeld in drie schalen (zie tabel 5).

52. De Vogel (2005).

53. Rogers (2000).

54. Hildebrand, De Ruiter & Van Beek (2001).

55. De Vogel, De Ruiter, Bouman & De Vries Robbé (2009).

Tabel 5: SAPROF items

<i>Interne factoren</i>	<i>Motivationale factoren</i>	<i>Externe factoren</i>
Intelligentie	Werk	Netwerk
Hechte band in de kindertijd	Vrijtijdsbesteding	Intieme relatie
Empathie	Financieel beheer	Hulpverlening
Coping	Motivatie voor behandeling	Woonsituatie
Zelfcontrole	Houding tegenover autoriteit	Toezicht
	Levensdoelen	
	Medicatie	

De factoren worden gescoord op een driepuntsschaal (0, 1, 2); 15 van de 17 items zijn dynamisch en daarmee veranderbaar door middel van interventie. De overige twee, Intelligentie en Hechte band in de kindertijd, zijn statisch. De SAPROF bevat de mogelijkheid tot het markeren van factoren die voor het betreffende individu het grootste belang hebben voor de bescherming. Dat kan door het aankruisen van 'key'-factoren: de belangrijkste aanwezige beschermende factoren voor het individu. Ook kunnen 'doel'-factoren worden aangemerkt; dat zijn factoren waar interventie zich als eerste op moet richten, omdat verbetering op die factoren naar verwachting extra bescherming oplevert. De SAPROF dient altijd in combinatie met het van toepassing zijnde risicotaxatie-instrument (B-SAFER, HCR-20^{V3}, HKT-R, RISC, SVR-20, et cetera) te worden gebruikt. Op basis van de SAPROF wordt het eindoordeel (op een driepuntsschaal variërend van weinig, redelijk tot veel) geformuleerd over de hoeveelheid aanwezige bescherming voor het individu in de situatie waarvoor de risicotaxatie wordt uitgevoerd. De laatste stap is de integratie van de SAPROF-uitslag met die van het gebruikte risicotaxatie-instrument, om een eindoordeel over het risico van het specifieke delictgedrag te vormen.

De risicotaxatie wordt gevolgd door een uitwerking van mogelijke risicoscenario's. Daarbij moeten de volgende vragen worden beantwoord: Wat zou er kunnen gebeuren? Wie zou slachtoffer kunnen worden? Hoe acuut en hoe ernstig zou het geweld zijn? Welke risico- en beschermende factoren zijn het meest van invloed op het geweldsrisico? De risicotaxatie wordt afgesloten met het formuleren van plannen voor risicomangement en interventies, die voortkomen uit de in de risicotaxatie in kaart gebrachte risico- en beschermende factoren, inclusief de *key*- en doel-factoren.

Inmiddels zijn er enige onderzoeken verricht naar de voorspellende waarde van de SAPROF en de meerwaarde daarvan ten opzichte van een taxatie die uitsluitend op risicofactoren is gericht. Die studies laten zien dat de dynamische factoren van de SAPROF de sterkste voorspellers zijn van non-recidive, zelfs bij een follow-up van elf jaar.⁵⁶ De voorspellende waarde voor de gecombineerde SAPROF-HCR-index was significant groter dan die voor de HCR alleen, en de SAPROF is voorspellend bij zowel gewelds- als seksuele delinquenten. Ook blijkt een toename van beschermende factoren gedurende een tbs-behandeling voorspellend te zijn voor minder recidive na afloop van de tbs-maatre-

56. De Vries Robbé, De Vogel & Douglas (2013).

gel.⁵⁷ Onderzoek uitgevoerd in Groot-Brittannië laat zien dat de SAPROF een significante voorspeller is van afwezigheid van geweld binnen de inrichting (AUC = .85).⁵⁸ Clinici zijn enthousiast over het gebruik van de SAPROF in de praktijk. Zij vinden het een waardevolle toevoeging die de risicotaxatie meer balans geeft, concrete handvatten biedt bij het formuleren van positieve behandeldoelen, en de communicatie met de delinquent over de risico's gemakkelijker maakt.

Risicocommunicatie

Het onderzoeksterrein van de communicatie over recidiverisico's is nog relatief onontgonnen. Effectieve risicocommunicatie helpt de besluitvormer te komen tot een heldere afweging van vrijheden en een eventuele terugkeer in de maatschappij. Zoals hierboven is geschetst, kan de literatuur over risicotaxatie worden onderverdeeld in twee benaderingen: het predictie- en het preventiemodel. Volgens het predictiemodel dient geweldsrisico te worden gecommuniceerd in termen van frequenties of waarschijnlijkheidspercentages op basis van actuariële risicotaxatie-instrumenten.⁵⁹ Het preventiemodel is een voorstander van risicocommunicatie in termen van categoriale schattingen (bijvoorbeeld laag, matig, hoog).⁶⁰ Daarbij dienen de voor de individuele delinquent relevante risicofactoren te worden beschreven, met nadruk op dynamische en beschermende factoren en een risicohanteringsplan dat het risico kan verminderen. Mills, Kroner en Morgan zijn voorstanders van een gemengd model van risicocommunicatie: (1) maak eerst een actuariële risicotaxatie op basis van groepsnormen; (2) identificeer vervolgens de belangrijke risicofactoren, met een speciale nadruk op dynamische factoren; en (3) identificeer interventie- en risicohanteringsstrategieën die het risico mogelijk kunnen verminderen.⁶¹

Voorals categoriale taxaties hebben nogal wat kritiek gekregen. Schopp⁶² maakte bezwaar tegen het categoriale format omdat het een element van subjectief oordeel bevat (wat als laag of hoog wordt gedefinieerd, beslist immers de ontwikkelaar of gebruiker van het instrument), waardoor de besluitvorming, die bij de rechter hoort te liggen, verschoven wordt naar de gedragsdeskundige. Bovendien bieden categoriale schattingen ook weinig stimulans aan de rechter om af te wijken van de risicobeoordeling die de gedragsdeskundige aanlevert.⁶³

Het omgekeerde geldt voor waarschijnlijkheidsschattingen, omdat de beoordelaar daarbij alleen beschrijvende gegevens levert en het oordeel over het risiconiveau aan de bevoegde wettelijke besluitvormer overlaat. Toch zijn ook waarschijnlijkheidsschattingen niet vrij van tekortkomingen. Ten eerste zijn er verschillende manieren waarop waarschijnlijkheidsschattingen over recidive kunnen worden gecommuniceerd. Het kan bijvoorbeeld zo: 'Meneer Z scoorde een XX op dit risicotaxatie-instrument. Van de personen met deze score recidiveert 40% binnen een follow-up van 5 jaar en 50% over 10 jaar.' Een alternatieve mogelijkheid is het presenteren van frequentieschattingen: 'Per-

57. De Vries Robbé, De Vogel, Douglas & Nijman (2015).

58. Abidin, Davoren, Naughton, Gibbons, Nulty & Kennedy (2013).

59. Bijvoorbeeld Harris & Rice (1984).

60. Douglas & Kropp (2002).

61. Mills, Kroner & Morgan (2011).

62. Schopp (1996).

63. Scurich & John (2012).

sonen met een score zoals meneer Z op dit risicotaxatie-instrument recidiveren met een kans van 4 op 10 over een follow-up van 5 jaar, en met 5 uit 10 over 10 jaar.’ Slovic, Monahan en MacGregor vergeleken deze twee risicocommunicatievormen en vonden dat het frequentie-format leidde tot een hogere risicoperceptie, zelfs als het risiconiveau, zoals in het voorbeeld hierboven, rekenkundig constant werd gehouden.⁶⁴ Gigerenzer toonde in diverse experimenten aan dat medici niet in staat zijn waarschijnlijkheidsschattingen adequaat te beoordelen.⁶⁵ Net als een rechter moet een medicus diagnostische testuitslagen beoordelen: was is de kans dat iemand de ziekte heeft, gegeven deze testuitslag? Wanneer de medici de testuitslagen als frequenties gepresenteerd kregen, begrepen zij de PPV veel beter. Onderzoek van Gigerenzer leidde ertoe dat het British Court of Appeal heeft aanbevolen dat DNA-bewijs voortaan in de vorm van frequenties wordt gepresenteerd. Dus: ‘Van elke 100.000 mensen zal er 1 een match laten zien’, in plaats van: ‘De kans op een match is 1 op 100.000.’

Recent toonden Scurich en John aan dat zelfs subtiele veranderingen in de wijze waarop waarschijnlijkheidsschattingen gepresenteerd worden, invloed hebben op de besluitvorming.⁶⁶ In een studie met 303 studenten die fungeerden als rechters, vonden zij dat risico geframed als een kans van 26% op geweld ertoe leidde dat de proefpersonen meestal besloten tot gedwongen opname in de psychiatrie, terwijl hetzelfde risico geframed als een kans van 74% op geen geweld, ertoe leidde dat de meeste studenten besloten tot vrijlating. Dit onderzoek laat zien dat door subtiele manipulatie van het frame van de risicoschatting, beslissingen over opsluiting en vrijlating krachtig kunnen worden gemanipuleerd.

De algemene trend in de praktijk van de risicocommunicatie om risico in termen van de kans op geweld te communiceren leidt er dus toe dat besluitvormers meer zullen neigen tot vals-positieve beslissingen. Hoewel Scurich en John studenten gebruikten als beslissers, is er geen reden om aan te nemen dat professionele rechters minder gevoelig zijn voor het effect van framing.⁶⁷ Scurich en John zijn van mening dat nader onderzoek moet vaststellen of het vermelden van de risicoschatting plus het complement daarvan (het risico van geen-geweld), het framing-effect tenietdoet. Als dat zo is, dan zouden gedragskundigen er verstandig aan doen in hun risicocommunicatie richting besluitvormers beide frequenties op te nemen.

Richtlijnen voor risicocommunicatie

Hoewel de empirische basis van de risicocommunicatie op het gebied van delinquent gedrag nog beperkt is, laat het huidige onderzoek wel zien dat zorgvuldige en evenwichtige communicatie over recidiverisico noodzakelijk is. In het licht daarvan suggereren wij een aantal voorlopige richtlijnen:

- Bij het rapporteren van waarschijnlijkheden afgeleid van actuariële risicotaxatie-instrumenten dient te worden vermeld dat die taxaties louter gebaseerd zijn op historische factoren en geen rekening houden met dynamische risicofactoren en protectieve factoren.

64. Slovic, Monahan & MacGregor (2000).

65. Gigerenzer (2002a).

66. Scurich & John (2011).

67. Guthrie, Rachlinski & Wistrich (2001).

- Rapporteer waarschijnlijkheidsschattingen niet alleen voor de negatieve uitkomst, maar ook voor de positieve uitkomst.⁶⁸
- Rapporteer in frequenties in plaats van in waarschijnlijkheidspercentages.
- Geef een overzicht van de belangrijkste dynamische risicofactoren en beschermende factoren in dit individuele geval en beschrijf of en hoe die factoren met elkaar interacteren.
- Geef expliciet aan of er in dit individuele geval nog casus-specifieke risico- of beschermende factoren van belang zijn die niet zijn opgenomen in het gebruikte risicotaxatie-instrument.
- Geef informatie over de context waarvoor de risicotaxatie geldig is (bijvoorbeeld in een gesloten setting, tijdens begeleid verlof, bij reclasseringstoezicht en zelfstandig wonen).
- Formuleer een risicohanteringsplan waarin staat aangegeven welke verschillende strategieën en interventies noodzakelijk zijn om het recidiverisico te verminderen.

Conclusies

Het afgelopen decennium is het gebruik van en het onderzoek naar gestructureerde risicotaxatie-instrumenten in de Nederlandse forensische sector enorm toegenomen. Op het eerste gezicht is dat een goede ontwikkeling, omdat daarmee de klinische blik van de jaren zeventig en tachtig een halt is toegevoerd. Bij nadere inspectie blijken ook aan de moderne, empirisch onderbouwde risicotaxatie-methoden risico's te zitten. De vergelijking met de chirurg en zijn gereedschap dringt zich op. Hij kan met zijn scalpel mensen genezen, maar hij kan er ook mensen mee beschadigen als hij de kennis en vaardigheden van zijn vak onvoldoende bijhoudt. Het 'vak' van de gestructureerde risicotaxatie is nog relatief jong, een stuk jonger dan de chirurgie, dus extra alertheid op verkeerd gebruik en onjuiste interpretatie van de uitkomsten is geboden.

Het psychologisch onderzoek bij Thomas in 2007 leidde tot de conclusie dat hij geen psychische stoornis had en een laag risico op recidive. Blijkbaar zette die contra-expertise de rechter die over de verlenging van zijn tbs moest beslissen aan het denken, want hij besloot Thomas voor observatie te laten opnemen in het Pieter Baan Centrum. De conclusies van het PBC-rapport uit 2008 komen sterk overeen met die van de contra-expertise. Thomas laat inzicht zien in de factoren die stress bij hem veroorzaken: een te hoog streefniveau, sub-assertiviteit, angst en dwangmatigheid. De rapporteurs zagen geen gronden die de longstay-status van Thomas konden rechtvaardigen, en men adviseerde beëindiging van de dwangverpleging en verlenging van de tbs met een jaar, met reclasseringstoezicht. Op 16 december 2008 beëindigde de rechter de dwangverpleging en Thomas verhuisde naar de woning van zijn zus.

Het laatste contact met Thomas en zijn zus dateert van februari 2015, in verband met een verzoek om toestemming om over zijn zaak te publiceren. Thomas woont zelfstandig. Vast werk vinden is hem niet meer gelukt sinds zijn terugkeer in de maatschappij, mede door het tbs-stigma en door de economische crisis. Nog steeds heeft Thomas regelmatig nachtmerries waarin personen en gebeurtenissen uit zijn jaren in de tbs-klinieken figureren. Hij is inmiddels zeven jaar uit de tbs, maar het lukt hem nog niet die periode uit zijn leven achter zich te laten.

68. Scurich & John (2012).

