

Voorspelling van recidive bij zedendelinquenten met behulp van retrospectief gebruik van de PCL-R en SVR-20

Koen Koster, Jacques van Lankveld, Marinus Spreen

van Mesdagkliniek, Groningen; Universiteit Maastricht

Samenvatting

In dit onderzoek zijn de SVR-20 en de PCL-R onderzocht op hun psychometrische kwaliteiten, wanneer beide instrumenten retrospectief gebruikt worden op basis van dossierinformatie. De predictieve validiteit is berekend over TBS-gestelden (N = 30) die at-risk waren. Daarnaast werden de recidivecijfers berekend over de at-risk populatie. De interbeoordelaarsbetrouwbaarheid en de interne consistentie zijn onderzocht door analyse van 58 dossiers van na 1984 uitgestroomde seksuele delinquenten van de Dr. S. van Mesdagkliniek.

De interbeoordelaarsbetrouwbaarheid was voor de PCL-R en de SVR-20 goed op instrumentniveau, op subschaalniveau redelijk tot goed. De subschaal 'seksuele delicten' van de SVR-20 scoorde slecht. Op itemniveau was de betrouwbaarheid acceptabel voor de niet-dynamische items van de PCL-R. Ook werd een acceptabele betrouwbaarheid gevonden voor de items van de SVR-20 behalve voor de items die behoorden tot de subschaal 'seksuele delicten'.

Voor de predictieve validiteit zijn door de kleine populatie en lage recidive base rates weinig significante resultaten gevonden voor seksuele recidive en gewelddadige, niet seksuele, recidive. Voor algemeen gewelddadige recidive werden significante voorspellers gevonden. De PCL-R en de factoren 'antisociale levensstijl' en 'levensstijl' van de PCL-R zijn goede voorspellers voor gewelddadige, niet seksuele, recidive en algemeen gewelddadige recidive. Van de SVR-20 voorspelt de totaalscore iets beter dan het gestructureerd klinisch oordeel. Beiden hebben een redelijke tot goede predictieve validiteit.

De base-rate voor seksuele recidive was 13%, voor gewelddadige, niet seksuele recidive, 23%. Voor algemeen gewelddadige recidive en enige recidive respectievelijk 33 en 50%.

De ontwikkeling van risicotaxatieinstrumenten voor de forensische gedragswetenschappen is van groot belang om de kans op toekomstige recidive zo nauwkeurig mogelijk in te kunnen schatten. Naast kennis over het inschatten van recidiverisico's levert onderzoek naar de prestaties van risicotaxatie instrumenten bruikbare informatie voor beleidsmakers op het gebied van de forensische psy-

chiatie. Ook kan empirisch onderzoek naar risicotaxatie en daaruit verkregen kennis bijdragen aan de ontwikkeling van theorievorming. Deze kennis en theorievorming kunnen naast handvatten voor bijvoorbeeld behandelbeleid een bijdrage leveren aan preventie-maatregelen (Bartels, Spreen, Bruinsma & Buschman, 2004).

In haar dissertatie heeft de Vogel (2005) de Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp & Webster, 1997; Nederlandse geautoriseerde vertaling: Hildebrand, de Ruiter & van Beek, 2001) geïntroduceerd in het Nederlands forensisch-gedragskundig onderzoek. Zij vond in haar studie een goede interbeoordelaarsbetrouwbaarheid voor de SVR-20. Daarnaast lijkt de SVR-20 eveneens niet-seksueel geweld goed te voorspellen. Als het gaat om seksuele recidive lijkt de gestructureerde klinische inschatting beter te presteren dan een zuiver actuaireel gebruik van het instrument. Strikt methodologisch gereedeneerd kunnen

drs. K. Koster, GZ-psycholoog i.o., Dr. S. van Mesdagkliniek, Postbus 30002, 9700 RC Groningen. E: k.koster@dj.minjus.nl

dr. J.J.D.M. van Lankveld, klinisch psycholoog, psychotherapeut, seksuoloog NVVS, Universiteit Maastricht, Departement Medische, Klinische en Experimentele psychologie.

dr. M. Spreen, socioloog en methodoloog, senior onderzoeker de Dr. S. van Mesdagkliniek, Groningen.

Dit artikel is gebaseerd op de doctoraalscriptie van de eerste auteur voor de opleiding Geestelijke Gezondheidskunde aan de Universiteit Maastricht.

Ontvangen: 8 maart 2006; Geaccepteerd: 21 augustus 2006.

de resultaten van het onderzoek van de Vogel niet zonder meer gegeneraliseerd worden naar de populatie van uitbehandelde zedendelinquenten. De onderzochte populatie van de Vogel is een selectieve steekproef omdat zij alleen tbs-gestelden bestudeert die een behandeling hebben ondergaan in de Dr. H. van der Hoevenkliniek. Om robuuste uitspraken te doen over de predictieve validiteit van de SVR-20 op landelijk niveau is een representatieve steekproef nodig waarin voor cultuur- en behandelverschillen tussen de verschillende tbs-klinieken gecontroleerd wordt. Daar deze er tijdens het schrijven van dit artikel niet was, hebben wij het onderzoek van de Vogel zo goed mogelijk gereproduceerd op de uitbehandelde populatie zedendelinquenten in de Mesdagkliniek. Onafhankelijke replicatie in andere klinieken zal bijdragen aan robuuste kennis van de SVR-20.

Naar de psychometrische kwaliteiten van de Psychopathy Checklist-Revised (PCL-R; Hare, 1991; Nederlandstalige versie: Vertommen, Verheul, de Ruiter & Hildebrand, 2002) is inmiddels veel (internationaal) onderzoek gedaan. In het meest recente Nederlandse onderzoek van het Expertisecentrum Forensische Psychiatrie (EFP) vonden Hildebrand, Hesper, Spreen en Nijman (2005) dat de PCL-R een goede voorspeller was voor gewelddadige recidive. Met name de factoren 'antisociale levensstijl' uit het eerste model van Hare en de factor 'levensstijl' uit het tweede model van Hare bleken goede predictoren (Hare, 1991; Hare 2003).

In dit onderzoek zijn de dossiers van alle 58 oud-patiënten die zijn veroordeeld voor een zedendelict, en uitstroonden tussen 1984 en 2003, van de Dr. S. van Mesdagkliniek gescoord op twee risicotaxatie instrumenten, namelijk de PCL-R en de SVR-20. De uitkomsten van de instrumenten zijn vergeleken met de recente justitiële gegevens van deze oud-patiënten waarbij is gekeken naar de mate waarin deze instrumenten recidiverisico kunnen taxeren.

Het doel van dit retrospectieve onderzoek is de kwaliteit en bruikbaarheid van risicotaxatie instrumenten voor zedendelinquenten te bestuderen en een bijdrage te leveren aan de ontwikkeling ervan.

De *probleemstelling* van dit onderzoek luidt als volgt: Wat is de predictieve validiteit van de SVR-20 en de PCL-R voor seksuele recidive en gewelddadige -niet seksuele- recidive bij (ex) ter beschikking gestelde seksuele delinquenten?

Daarnaast is er gekeken naar de interbeoordelaarsbetrouwbaarheid en interne consistentie.

Methoden van onderzoek

Steekproef

De steekproef bestond uit alle 58 mannelijke seksuele delinquenten die minimaal twee jaar in de Dr. S. van Mesdagkliniek hebben verbleven in de periode van 1984-2002.

De volgende inclusie- en exclusiecriteria zijn van toepassing:

- Het index delict is conform art. 240-249 Sr. of een verkrachtingsdelict in de justitiële voorgeschiedenis
- IQ > 70
- Dossier is van voldoende kwaliteit om te kunnen scoren

Procedure

Het onderzoek heeft een retrospectief follow-up design. Op basis van de dossiers van (ex) ter beschikking gestelden (bestaande uit processen-verbaal, Pro Justitia rapportages, een (ziekte)geschiedenis, behandelplannen en -evaluaties en verlengingsadviezen) zijn de PCL-R en de SVR-20 gescoord door twee beoordelaars. Uit deze dubbelscore komt een consensuscore voort welke is gebruikt in de berekeningen. Acht beoordelaars hebben meegewerkt aan het onderzoek, allen waren zij getraind in het gebruik van de PCL-R en de SVR-20.

De recidivegegevens zijn via het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) verkregen bij de Centrale Justitiële Documentatiedienst (CJD) te Almelo.

De recidivegegevens zijn opgevraagd nadat alle 58 dossiers gescoord waren.

Scoringsprotocol

Alle 58 dossiers zijn onafhankelijk van elkaar dubbel gescoord door twee beoordelaars. Op basis van deze dubbele scoring komt een consensuscore tot stand die is gebruikt voor de analyses van de predictieve validiteit. Daarnaast werden de onafhankelijke scores gebruikt voor het bepalen van de interbeoordelaarsbetrouwbaarheid.

Recidivegegevens

De recidivegegevens zijn opgevraagd bij het Justitieel Documentatieregister van het Ministerie van Justitie. Recidive is gedefinieerd als:

Een nieuwe veroordeling voor een delict tijdens de follow-up periode, zoals omschreven in het Wetboek van Strafrecht. Vier categorieën recidive worden onderscheiden.

- A Seksuele recidive is gedefinieerd als een nieuwe veroordeling voor een seksueel delict volgens het Wetboek van Strafrecht.
- B Gewelddadige niet-seksuele recidive verwijst naar een nieuwe veroordeling voor niet-seksuele gewelddadige delicten, zoals (poging tot) mishandeling, doodslag, brandstichting.
- C Algemeen gewelddadige recidive betreft een nieuwe veroordeling voor een gewelddadig niet-seksueel delict of een seksueel delict. (A en/of B)
- D Enige recidive verwijst naar een veroordeling voor overige delicten bijvoorbeeld vermogens, verkeers- en drugsdelicten. Delicten die vallen onder categorie A, B en C worden niet in categorie D verrekend.

Dataverwerking en analyse

Voor de vergelijkbaarheid met de Vogel (2005) zal de interbeoordelaarsbetrouwbaarheid worden gemeten met behulp van de Intraklasse Correlatie Coëfficiënt (ICC), waarbij gebruik wordt gemaakt van het two-way random effects variantie model en consistency type (McGraw & Wong, 1996; Hildebrand, 2004). Zowel de *single measure* als *average measure* ICC's worden gemeten.

De predictieve validiteit van de risicotaxatie instrumenten als ook van de specifieke items zullen worden onderzocht door middel van Receiver Operating Characteristics (ROC) analyses, uitgedrukt als Area Under the Curve (AUC)-scores. ROC-analyse is een statistische methode die internationaal gebruikt wordt als toets voor de voorspellende waarde van een variabele of instrument. De methode bestaat hieruit dat de percentages juiste en onjuiste voorspellingen, respectievelijk de percentages true en false positives grafisch geplot worden door middel van een curve. De oppervlakte van het gebied onder deze curve tot aan de rechte lijn die het 0-punt verbindt met de top rechtsboven in de figuur is de AUC-score. De rechte lijn verwijst naar toevalspredictie (het aantal juiste voorspellingen is gelijk aan het aantal onjuiste voorspellingen) en heeft een AUC-waarde van .50. Een AUC-waarde kan als volgt geïnterpreteerd worden (Philippe, 2005): de kans dat een toevallig gekozen recidivist een hogere score op het instrument heeft dan een toevallig niet gekozen recidivist. Het voordeel van de ROC-methode is de onafhankelijkheid van de base rate van de uitkomstmaat. Voor een gedetailleerdere uitleg over ROC-analyse verwijzen wij naar Mossman (1994).

Resultaten

Demografische gegevens

Van de 58 ex-patiënten die sinds 1984 de kliniek verlaten hebben zijn 33 overgeplaatst naar een andere kliniek, 25 kregen ontslag vanuit de Dr. S. van Mesdagkliniek door een beëindiging van de TBS-maatregel. Van de 33 overgeplaatste patiënten hebben vijf inmiddels een beëindiging van de TBS-maatregel gekregen. In totaal zijn er dus 30 ex-patiënten at-risk geweest, dat wil zeggen hebben in de maatschappij gefunctioneerd zonder een TBS-maatregel. De at-risk periode is zuiver gebaseerd op de periode vanaf de opheffing van de TBS-maatregel, onbegeleide verloven en proefverlof zijn niet meegenomen in de at-risk tijd. Het was onmogelijk om voor iedereen terug te vinden hoe lang (ex)patiënten onbegeleid buiten de kliniek hebben doorgebracht ten tijde van de TBS-maatregel in de vorm van onbegeleid verlof of proefverlof. 28 patiënten vallen ten tijde van dit onderzoek onder een TBS-maatregel. Het gemiddelde aantal opnamemaanden in de Mesdagkliniek was voor de gehele populatie 75 maanden, voor de groep bij wie de TBS beëindigd was lag het gemiddelde aantal opnamemaanden op 71. De etnische achtergrond van de gehele populatie was voornamelijk Nederlands (78%), 9 % had een Surinaamse afkomst, 3% van de populatie was Antilliaans, 10 % had een andere afkomst. De gemiddelde at-risk tijd voor de at-risk populatie was 106 maanden (SD: 57, range 16-252 maanden).

Interbeoordelaarsbetrouwbaarheid en interne consistentie

De betrouwbaarheid van de PCL-R en de SVR-20 is samengevat in tabellen 1 en 2. Bij de PCL-R is vooral minder overeenstemming op de meer dynamische items. Op schaal- en subschaalniveau worden voor de PCL-R acceptabele tot uitstekende interbeoordelaarsbetrouwbaarheidsscores gehaald met single rater ICC-waarden tussen de .54 en .81. De interne consistentie van de PCL-R en de subschaal 'levensstijl' is goed met betrouwbaarheidcoëfficiënten van respectievelijk Cronbach's $\alpha = .79$ en $\alpha = .76$. De factoren uit het tweefactoren model scoren redelijk met Cronbach's $\alpha = .69$ per factor. De subschalen 'affectief' en 'antisociaal' uit het vierfactoren model hebben een lage interne consistentie met alfa's van respectievelijk .58 en .53¹.

De items van de SVR-20 die slecht scoren zijn met name te vinden in de subschaal 'seksuele delicten'.

¹ Het tweefactorenmodel van Hare (1991) bestaat uit de factoren: 1. agressief narcisme en 2. antisociale levensstijl. Het vierfactorenmodel van Hare (2003) bestaat uit de factoren 1. interpersoonlijk; 2. affectief; 3. levensstijl; 4. antisociaal.

Tabel 1 Interbeoordelaarsbetrouwbaarheid en interne consistentie van de PCL-R (N = 58)

Item/schaal	Single Rater ICC	Cronbach's α
1: glatte prater/ oppervlakkige charme	.50***	
2: sterk opgeblazen gevoel van eigenwaarde	.65***	
3: prikkelhongerig/ neiging tot verveling	.35*	
4: pathologisch liegen	.38**	
5: list en bedrog/ manipulerend gedrag	.59***	
6: gebrek aan berouw of schuldgevoel	.35**	
7: ontbreken van emotionele diepgang	.24	
8: kil/ gebrek aan empathie	.15	
9: parasitaire levensstijl	.38**	
10: gebrekkige beheersing van het gedrag	.52***	
11: promiscue seksueel gedrag	.22	
12: gedragsproblemen op jonge leeftijd	.56***	
13: ontbreken van realistische doelen op lange termijn	.43***	
14: impulsiviteit	.27	
15: onverantwoordelijk gedrag	.36**	
16: geen verantwoordelijkheid nemen voor het eigen gedrag	.23	
17: veel kortstondige partnerrelaties	.42**	
18: jeugdcriminaliteit	.61***	
19: schending van voorwaarden bij voorwaardelijke veroordeling en/ of achterwege blijven van vervroegde of voorwaardelijke invrijheidsstelling	.47***	
20: veelsoortige criminaliteit	.45***	
PCL-R totaalscore	.81***	.79
<u>Tweefactorenmodel:</u>		
PCL-R factor I: agressief narcisme	.67***	.69
PCL-R factor II: antisociale levensstijl	.75***	.69
<u>Vierfactorenmodel:</u>		
PCL-R factor I: interpersoonlijk (items: 1,2,4,5)	.70***	.64
PCL-R factor II: affectief (items: 6,7,8,16)	.54***	.58
PCL-R factor III: levensstijl (items: 3,9,13,14,15,)	.63***	.76
PCL-R factor IV: antisociaal (items: 10, 12,18,19,20)	.70***	.53

Noot: * = $p < .01$, ** = $p < .005$, *** = $p < .001$

Op schaal en subschaal niveau voor de SVR-20 worden acceptabele tot goede interbeoordelaarsbetrouwbaarheidsscores gehaald met single rater ICC waarden tussen de .42 en .72.

Van de SVR-20 is de interne consistentie van de gehele schaal acceptabel (Cronbach's $\alpha = .72$). De subschalen 'sociale aanpassing' en 'seksuele delicten' zijn weinig consistent met alfa's van resp. .51 en .54. De schaal 'toekomstplannen' lijkt daarentegen een intern zeer consistente schaal met een alfa van .84.

Recidive

In totaal recidiveerden 19 van de 30 ex-patiënten voor enig delict (63%) met een gemiddelde follow-up periode van 106 maanden (SD: 57, range: 16-252). Van

de overgeplaatste vijf personen waarvan de maatregel elders werd beëindigd pleegden alle vijf enige vorm van recidive tegen 56% van hen wiens maatregel werd beëindigd gelijktijdig met hun uitstroom uit de Mesdagkliniek. De negentien recidivisten waren gezamenlijk verantwoordelijk voor 96 nieuwe veroordelingen wat neerkomt op een gemiddelde van vijf nieuwe veroordelingen per recidivist (zie tabel 3).

Dertien procent van de at-risk populatie pleegde opnieuw een seksueel delict. Na beëindiging maatregel pleegde 23% een ernstig (niet seksueel) agressief delict. 50% recidiveerde met een lichter vergrijp (zie tabel 4).

Geen van de plegers van seksuele recidive pleegde meer dan een seksuele recidive per persoon. Zij die

Tabel 2. Interbeoordelaarsbetrouwbaarheid en interne consistentie van de SVR-20 (N = 58)

Item/schaal	Single Rater ICC	Cronbach's α
1: seksuele deviatie	.42**	-
2: slachtoffer van kindermishandeling	.60***	
3: psychopathie	.64***	
4: ernstige psychiatrische stoornis	.45***	
5: problemen met het gebruik van middelen	.51***	
6: suïcidale/homocidale gedachten	.41**	
7: relatieproblemen	-.05	
8: problemen met betrekking tot het arbeidsverleden	.33*	
9: eerder gepleegde, niet-seksueel gewelddadige, delicten	.42**	
10: eerder gepleegde, niet-seksueel gewelddadige, niet-seksuele, delicten	.58***	
11: eerdere schending van voorwaarden	.52***	
12: hoge dichtheid van seksuele delicten	.56***	
13: meerdere typen seksuele delicten	.14	
14: lichamelijk letsel bij slachtoffer(s) van seksuele delicten	.45***	
15: gebruik van wapens of bedreigingen met de dood tijdens plegen van seksuele delicten	.32	
16: escalatie in frequentie of ernst van seksuele delicten	.23	
17: extreme minimalisering of ontkenning van seksuele delicten	.35*	
18: opvattingen waaruit goedkeuring of vergoelijking van seksuele delicten blijkt	.34*	
19: ontbreken van realistische toekomstplannen	.49***	
20: negatieve houding ten opzichte van interventies	.45***	
SVR totaalscore	.72***	.72
SVR schaal I: sociale aanpassing	.63***	.51
SVR schaal II: seksuele delicten	.42**	.54
SVR schaal III: toekomstplannen	.61***	.84

Noot: * = $p < .01$, ** = $p < .005$, *** = $p < .001$

recidiveerden voor een geweldsdelict waren gemiddeld verantwoordelijk voor twee nieuwe veroordelingen. Voor enige recidive kregen 15 personen 79 nieuwe veroordelingen, waarbij vermeld moet worden dat twee personen voor respectievelijk 25 en 13 nieuwe veroordelingen verantwoordelijk waren.

Predictieve validiteit

In tabel 5a, 5b en 5c zijn de AUC waarden van de PCL-R en de SVR-20 samengevat voor respectievelijk seksuele recidive, gewelddadige maar niet seksuele recidive en algemeen gewelddadige recidive (seksueel en niet-seksueel). Voor seksuele recidive zijn geen significante AUC waarden gevonden ($p > .05$). Voor gewelddadige, niet seksuele, recidive voorspelt de totaalscore voor de PCL-R goed met een AUC waarde van .77 ($p < .05$). De factor 'antisociale le-

vensstijl' uit het tweefactoren model doet het nog iets beter met een AUC waarde van .79 ($p < .05$). Bij het voorspellen van algemeen gewelddadige recidive is het merendeel van de verkregen uitkomsten significant. Van de PCL-R zijn de totaalscore, de factor 'antisociale levensstijl' uit het tweefactoren model en de factor 'levensstijl' uit het vierfactoren model goede voorspelers voor algemeen gewelddadige recidive. Bij de SVR-20 voorspelt de totaalscore het best en iets beter dan het gestructureerd klinisch oordeel. Het verschil in prestatie tussen de (sub) schalen met significante uitkomsten is klein.

Discussie

In dit retrospectieve follow-up onderzoek zijn de betrouwbaarheid en de predictieve validiteit van de PCL-R en de SVR-20 onderzocht onder de populatie van

Tabel 3. Recidive naar aantal veroordelingen (N = 91) per delicttype bij een at-risk populatie van N = 30.

	Aantal veroordelingen	percentage van het totaal aantal veroordelingen
A seksuele recidive	4	4.2 %
B gewelddadige, niet seksuele, recidive	13	13.5 %
D enige recidive	79	82.3 %
Totaal	96	

Tabel 4. Recidivecijfers at-risk populatie per delicttype N = 30

	Aantal recidivisten	recidivepercentage
A seksuele recidive	4	13.3 %
B gewelddadige, niet seksuele, recidive	7	23.3 %
C algemeen gewelddadige recidive*	10	33.3 %
D enige recidive	15	50.0 %

Noot: * een persoon werd na beëindiging maatregel veroordeeld voor verkrachting, en later nogmaals veroordeeld voor moord.

alle na 1984 uitgestroomde patiënten van de Dr. S. van Mesdagkliniek die veroordeeld waren voor een seksueel delict. De betrouwbaarheid en interne consistentie zijn berekend over de totale populatie van 58 personen. De berekeningen voor de predictieve validiteit zijn gedaan over hen die at-risk waren (N = 30).

Interbeoordelaarsbetrouwbaarheid en interne consistentie

De mate van overeenstemming bij de PCL-R was uitstekend op schaalniveau. Op factorniveau was de overeenstemming voor het tweefactorenmodel redelijk tot goed. In het vierfactorenmodel scoorde alleen factor twee 'affectief' matig, de overige factoren scoorden redelijk tot goed. Hildebrand en de Ruiter (2002) vonden op schaal- en factorniveau hogere waarden. De onderlinge relatieve verhouding tussen de schaal en factorscores (op het tweefactorenmodel), gevonden in het huidige onderzoek, is qua structuur vergelijkbaar.

Op itemniveau werden lage scores voor interbeoordelaarsbetrouwbaarheid gevonden voor tien items. Een verklaring voor de slechte overeenstemming op deze items zou kunnen zijn dat in dit onderzoek geen gebruik gemaakt kon worden van een klinisch interview. Voor bovenstaande items, veelal met

een dynamisch karakter, maakt dit het nauwkeurig scoren moeilijker. Op basis van alleen dossierinformatie kan gekozen worden voor verschillende interpretaties omdat het dossier over een langere tijd verslag doet van dynamische en dus variabele toestanden. Ook een lagere kwaliteit van dossiers kan leiden tot lagere ICC-waarden. De interne consistentie voor de PCL-R als geheel was goed. Ook de factor 'levensstijl' bleek een consistente schaal te zijn. Hildebrand en de Ruiter (2002) vonden in hun onderzoek wederom hogere maar als het gaat om onderlinge relatieve verhouding tussen de schaal en het tweefactorenmodel vergelijkbare scores voor interne consistentie. De interbeoordelaarsbetrouwbaarheid van de SVR-20 was goed. Redelijke ICC waarden werden gemeten voor de schalen 'sociale aanpassing' en 'toekomstplannen'. De overeenstemming op de subschaal 'seksuele delicten' was zeer matig. Een verklaring voor deze lagere score is moeilijk te geven. Juist deze informatie is doorgaans betrouwbaar te scoren uit dossierinformatie. Het zou kunnen zijn, met name bij de oudere dossiers, dat de dossierinformatie te wensen over heeft gelaten.

De interne consistentie van de SVR-20 was goed, ook de schaal 'toekomstplannen' had een uitstekende interne consistentie. De eerste twee schalen (sociale aanpassing, seksuele delicten) waren intern weinig

consistent. De Vogel (2005) vond in haar dissertatie-onderzoek vergelijkbare interbeoordelaarsbetrouwbaarheidscoëfficiënten en een hogere mate van overeenstemming voor de drie subschalen.

Samenvattend kan over de interne consistentie en de interbeoordelaarsbetrouwbaarheid van de PCL-R en de SVR-20 het volgende gezegd worden. Zowel de PCL-R als, in iets mindere mate, de SVR-20 hebben een goede mate van interbeoordelaarsbetrouwbaarheid op schaalniveau. De schalen zijn intern vrij

consistent. Op subschaalniveau zijn de psychometrische kwaliteiten wisselend maar gemiddeld als redelijk te kwalificeren. Op itemniveau zijn de resultaten dermate dat items van de twee instrumenten (enkele uitzonderingen daargelaten) vooralsnog ongeschikt zijn voor klinisch gebruik. Dat betekent dat de toepassing van dynamische items bij evaluatie van behandeling, bijvoorbeeld, op dit moment niet verantwoord is. Dat geldt in ieder geval als er geen klinische indrukken zijn op basis waarvan de betrouwbaarheid mogelijk omhoog gaat.

Tabel 5a. AUC waarden voor seksuele recidive

	AUC	95% betrouwbaarheidsinterval
PCL-R totaalscore	.71	.49 - .93
<u>Tweefactorenmodel:</u>		
PCL-R factor I: agressief narcisme	.67	.46 - .88
PCL-R factor II: antisociale levensstijl	.69	.50 - .88
<u>Vierfactorenmodel:</u>		
PCL-R factor I: interpersoonlijk	.76	.59 - .94
PCL-R factor II: affectief	.44	.14 - .75
PCL-R factor III: levensstijl	.71	.53 - .89
PCL-R factor IV: antisociaal	.56	.27 - .85
SVR totaalscore	.63	.37 - .88
SVR gestructureerd klinisch oordeel	.69	.47 - .91
SVR schaal I: sociale aanpassing	.62	.33 - .91
SVR schaal II: seksuele delicten	.54	.23 - .85
SVR schaal III: toekomstplannen	.64	.39 - .88

Noot: geen enkele AUC waarde was significant op het niveau $p < .05$

Tabel 5b. AUC waarden voor geweldadige, niet seksuele, recidive

	AUC	95% betrouwbaarheidsinterval
PCL-R totaalscore	.77*	.59 - .94
<u>Tweefactorenmodel:</u>		
PCL-R factor I: agressief narcisme	.66	.44 - .88
PCL-R factor II: antisociale levensstijl	.79*	.61 - .97
<u>Vierfactorenmodel:</u>		
PCL-R factor I: interpersoonlijk	.65	.44 - .86
PCL-R factor II: affectief	.64	.41 - .86
PCL-R factor III: levensstijl	.73	.53 - .93
PCL-R factor IV: antisociaal	.74	.55 - .93
SVR totaalscore	.74	.55 - .92
SVR gestructureerd klinisch oordeel	.72	.53 - .90
SVR schaal I: sociale aanpassing	.71	.52 - .89
SVR schaal II: seksuele delicten	.67	.48 - .86
SVR schaal III: toekomstplannen	.69	.48 - .89

Noot: * = $p < .05$

Tabel 5c. AUC waarden voor algemeen geweldadige recidive

	AUC	95% betrouwbaarheidsinterval
PCL-R totaalscore	.82**	.66 - .97
<u>Twedefactorenmodel:</u>		
PCL-R factor I: agressief narcisme	.69	.50 - .88
PCL-R factor II: antisociale levensstijl	.83***	.68 - .97
<u>Vierfactorenmodel:</u>		
PCL-R factor I: interpersoonlijk	.74*	.56 - .92
PCL-R factor II: affectief	.55	.33 - .77
PCL-R factor III: levensstijl	.80**	.64 - .95
PCL-R factor IV: antisociaal	.73*	.54 - .92
SVR totaalscore	.78*	.62 - .94
SVR gestructureerd klinisch oordeel	.75*	.58 - .92
SVR schaal I: sociale aanpassing	.75*	.57 - .92
SVR schaal II: seksuele delicten	.66	.46 - .86
SVR schaal III: toekomstplannen	.73*	.55 - .91

Noot: * = $p < .05$; ** = $p < .01$; *** = $p < .005$

Tabel 6. De voorspellende waarde van totaalscore en subschaalscores van de PCL-R en de SVR-20 voor geweldadige niet-seksuele recidive en algemeen geweldadige recidive van TBS-behandelenden in Nederlandse studies

	Gevonden waarden van de Area Under the Curve							
	Gewelddadige niet seksuele recidive				Algemene geweldadige recidive			
	A	B	D	E	A	B	C	E
PCL-R totaalscore	.77*	.66*		.72***	.82**	.70**	.75***	.69***
<u>Twedefactoren model</u>								
PCL-R factor I	.66	.55		.61	.69	.62	.63**	.57
PCL-R factor II	.79*	.68**		.77***	.83***	.69**	.79***	.74***
<u>Vierfactoren model</u>								
PCL-R factor I	.65			.57	.74*		.55	.56
PCL-R factor II	.64			.62*	.55		.67**	.58
PCL-R factor III	.73			.77***	.80**		.77***	.77***
PCL-R factor IV	.74			.74***	.73*		.77***	.76***
SVR totaalscore	.74		.71***		.78*			
SVR gestructureerd klinisch oordeel	.72		.67**		.75*			
SVR sociale aanpassing	.71		.71***		.75*			
SVR seksuele delicten	.67		.61		.66			
SVR toekomstplannen	.69		.70***		.73*			

Noot: * = $p < .05$; ** = $p < .01$; *** = $p < .005$

A: Koster, van Lankveld & Spreen (2006)

B: Hildebrand, de Ruiter, & de Vogel (2003)

C: de Vogel, de Ruiter, Hildebrand, Bos & van de Ven (2004)

D: de Vogel (2005)

E: Hildebrand, Hesper, Spreen & Nijman (2005)

Hier is dus een verschil tussen theoretisch mogelijke bruikbaarheid, die groot is, en de huidige praktische mogelijkheden refererend aan de matige psychometrische resultaten op itemniveau.

De gevonden waarden zijn, hoewel gemiddeld lager, vergelijkbaar met het nog summier Nederlandse onderzoek als het gaat om onderlinge relatieve verhouding van de (sub) schalen. Een verklaring voor de gemiddeld lagere gevonden waarden zou de geringere steekproefomvang kunnen zijn. Een andere verklaring kan de pluriformere en grotere groep van beoordelaars (psychologen, psychiaters, en onderzoekers) zijn in dit onderzoek, i.t.t. ander onderzoek waar met een beperkter aantal beoordelaars met eenzelfde functie (onderzoeker) gewerkt werd.

Recidive

De recidivecijfers voor de onderzochte populatie zijn 13% voor seksuele recidive, 23% voor gewelddadige, niet seksuele, recidive, 33% voor algemeen gewelddadige recidive en 50% voor enige recidive. Zoals eerder vermeld moeten deze cijfers worden beschouwd als ondergrens omdat niet alle delicten tot een strafrechtelijke afhandeling komen. De gevonden recidivecijfers komen overeen met ander recent onderzoek. Wartna en collega's (2005) vinden een recidive cijfer van 11% voor seksuele recidive bij tbs-gestelde seksueel delinquenten over een follow-up periode van 120 maanden. Voor geweldsrecidive bij ter beschikking gestelde geweldsdelinquenten vonden zij een recidive van 27% met eenzelfde follow-up periode. Philipse (2005) vond een recidivecijfer van 15% voor seksuele recidive en 18% voor gewelddadige recidive bij een gemiddelde follow-up periode van 62 maand. De Ruiters en de Vogel (2004) vinden wel hogere base rates. Voor seksuele recidive lag dit op 39%, voor gewelddadige niet seksuele recidive op 46%. Hun gemiddelde follow-up periode was 140 maanden.

Predictieve validiteit

De resultaten aangaande de predictieve validiteit van de PCL-R en de SVR-20 waren voor de categorie 'seksuele recidive' in zijn geheel niet significant, voor de categorie 'gewelddadige, niet seksuele, recidive' zijn een aantal significante resultaten gevonden, terwijl de totaalscores en subschalen van PCL-R en SVR-20 bijna allemaal goede voorspellers bleken te zijn van 'algemeen gewelddadige recidive'. De oorzaken van de beperkte bevindingen in de eerste twee recidivecategorieën zijn waarschijnlijk de kleine omvang van de populatie en de lage base-rate. Vooral de kleine populatie is een duidelijke beperking van dit onderzoek. Hierdoor kon er ook geen onderscheid ge-

maakt worden tussen verschillende typen plegers van seksueel geweld zoals incestplegers, verkrachters en pedoseksuelen in dit onderzoek. Dat is jammer want als het gaat om bijvoorbeeld recidive vallen incestplegers minder vaak terug dan verkrachters en zijn seksuele delinquenten dus niet als een homogene groep te beschouwen (de Ruiters & de Vogel, 2004). Omdat alleen uitspraken gedaan kunnen worden over statistisch significante resultaten zullen alleen deze worden besproken in deze discussie.

De predictieve validiteit van de PCL-R totaalscore is goed voor gewelddadige, niet seksuele-, en algemeen gewelddadige recidive. De factor 'antisociale levensstijl' uit het tweefactoren model lijkt nog beter te voorspellen voor deze twee recidivecategorieën. Uit het vierfactoren model komt de factor 'levensstijl' naar voren als beste voorspeller voor 'algemeen gewelddadige recidive'. De factoren 'interpersoonlijk' en 'antisociaal' zijn redelijk voorspellend voor 'algemeen gewelddadige recidive'.

Van de SVR-20 voorspelt de totaalscore van de schaal algemeen gewelddadige recidive het beste. De totaalscore voorspelt daarmee beter dan het gestructureerd klinisch oordeel. De subschalen 'sociale aanpassing' en 'toekomstplannen' hebben een redelijke voorspellende waarde voor 'algemeen gewelddadige recidive'.

Andere Nederlandse onderzoeken (Hildebrand e.a., 2003, de Vogel e.a., 2004) vonden vergelijkbare resultaten voor de PCL-R. Op subschaal-niveau bleken de tweede factor 'antisociale levensstijl' uit het tweefactoren model en de factor 'levensstijl' uit het vierfactoren model goede voorspellers.

De Vogel (2005) verkreeg lagere AUC waarden voor de SVR-20 bij algemeen gewelddadige recidive maar in haar onderzoek voorspelde de totaalscore bij de besproken recidivecategorieën eveneens beter dan het gestructureerd klinisch oordeel. Voor een vergelijkend overzicht van onderzoeksuitkomsten zie tabel 6.

Conclusie

De vraagstelling van dit onderzoek luidde: "Wat is de predictieve validiteit van de PCL-R en de SVR-20 voor het voorspellen van (seksuele) recidive bij ex ter beschikkinggestelde seksuele delinquenten?" De totaalscore van de PCL-R heeft een behoorlijke predictieve validiteit ten aanzien van gewelddadige, niet seksuele, recidive en algemeen gewelddadige recidive. De factor 'antisociale levensstijl' uit het tweefactoren model presteert nog iets beter dan de gehele schaal. Ook de factor 'levensstijl' uit het vierfactoren model van psychopathie volgens Hare (2003) valt op door goede prestaties: een redelijk goede interbeoordelaars-

betrouwbaarheid, een behoorlijk goede interne consistentie, en een goede voorspelling van algemeen gewelddadige recidive. Dit betekent dat een criminele historie en de bijbehorende levensstijl goede voorspelers zijn van een voortzetting van die levensstijl. De SVR-20 presteert op totaalscoreniveau beter dan op het niveau van het gestructureerd klinisch oordeel, maar beiden zijn een redelijke voorspeller voor algemeen gewelddadige recidive. De subschaal 'seksuele delicten' valt op door een lage betrouwbaarheid en een slecht voorspellend vermogen.

Alhoewel de totaalscore in dit onderzoek beter voorspelde dan het gestructureerd klinisch oordeel, moet men zich niet blind staren op totaalscores. De SVR-20 is bedoeld voor gestructureerde klinische taxatie: individuele factoren en weging van de afzonderlijke items zouden een rol moeten spelen. Het instrument draagt bij aan de behoefte om risico-inschattingen van deskundigen meer inzichtelijk te maken voor leken. Huidig onderzoek bevestigt de meerwaarde van de PCL-R en de SVR-20 voor de inschatting van toekomstig geweld, en de resultaten komen grotendeels overeen met de resultaten die tot op heden gevonden zijn in het Nederlandse onderzoek naar de voorspellende waarde van risicotaxatie.

Voor vervolgonderzoek verdient het de aanbeveling om met een grotere populatie te werken. Door de grotere statistische power die hierdoor verkregen wordt kunnen 'ware' significante resultaten beter worden aangetoond. Bovendien zoals eerder gesteld: om robuuste uitspraken te doen over de predictieve validiteit van de SVR-20 op landelijk niveau is een representatieve steekproef nodig waarin voor cultuur- en behandelverschillen tussen de verschillende tbs-klinieken gecontroleerd wordt

Summary

Risk assessment among sexual delinquents using PCL-R and SVR-20 data from retrospective record analysis

In this study, the interrater reliability and internal consistency of the SVR-20 and PCL-R were studied in a sample of sexual delinquents who were treated in and released from a forensic psychiatric hospital in The Netherlands. The predictive validity was assessed in the at risk population (N = 30).

The interrater reliability of the PCL-R and SVR-20 was good at the scale level, and average to good at the subscale level. The subscale 'sexual offences' of the SVR-20 was poor in this respect. At item level the reliability was average for the non-dynamic items of the PCL-R. An acceptable reliability was also found of the items of the SVR-20, except for the

sub-scale 'sexual offences'.

Prediction of sexual recidivism and violent, non sexual, recidivism was poor, probably due to the small sample and the low base-rate levels for reconviction. Better prediction was found for general violent recidivism. The PCL-R total scale and the subscales 'antisocial lifestyle' (according to the two-factor model of psychopathy, Hare, 1991) and 'lifestyle' (according to the four-factor model of psychopathy, Hare, 2003) are good predictors of violent non-sexual recidivism and general recidivism. The total score of the SVR-20 predicts somewhat better than the structured clinical judgment, but both have moderate predictive validity. For sexual recidivism and violent, non sexual, recidivism base rates were found of 13 and 23%. For general recidivism and any recidivism base rates were examined of 33 and 50%.

Literatuur

- Bartels, A.A.J., Spren, M., Bruinsma, C.L., & Buschman, J. (2004). *Risicotaxatie in de Dr. S. van Mesdagkliniek, versie 5*. Interne publicatie. Dr. S. van Mesdagkliniek: Groningen.
- Boer, D.P., Hart, S.D., Kropp, P.R., & Webster, C.D. (1997). *Manual for the Sexual Violence Risk-20. Professional guidelines for assessing risk of sexual violence*. Vancouver, British Columbia: Institute against Family Violence.
- Hare, R.D. (1991). *The Hare Psychopathy Checklist-Revised*. Toronto, Ontario: Multi-Health Systems.
- Hare, R.D. (2003). *Hare Psychopathy Checklist-Revised (PCL-R): 2nd edition*. Toronto, Canada. Multi-Health Systems.
- Hildebrand, M., de Ruiter, C., & van Beek, D. (2001). SVR-20. *Richtlijnen voor het beoordelen van het risico van seksueel gewelddadig gedrag*. Utrecht: Dr. Henri van der Hoeven Stichting.
- Hildebrand, M., & de Ruiter, C. (2002). Kort instrumenteel. De Nederlandstalige versie van Hare's Psychopathy Checklist-Revised: enige psychometrische bevindingen. *Gedragstherapie*, 35, 329-339.
- Hildebrand, M., de Ruiter, C., & de Vogel, V. (2003). Recidive van verkrachters en aanranders na tbs: de relatie met psychopathie en seksuele deviatie. *De Psycholoog*, 38, 114-124.
- Hildebrand, M., Hesper, B.L., Spren, M., & Nijman, H.L.I. (2005). De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie. EFP. Actuele Kennis 1.
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62, 783-792.
- Philipse, M.W.G. (2005). *Predicting criminal recidivism. Empirical studies and clinical practice in forensic psychiatry*. Nijmegen: Radboud Universiteit [dissertatie].
- Philipse, M.W.G. (2005). *Inschatten van het risico van delictherhaling in de TBS. Resultaten uit een longitudinale studie. Symposium Succes in de TBS??!! – ggze*, 28 april 2005.

- Ruiter de, C., & de Vogel, V. (2004). Recidive bij behandelde seksuele delinquenten. *Tijdschrift voor Seksuologie*, 28, 92-102.
- Vertommen, H., Verheul, R., de Ruiter, C., & Hildebrand, M. (2002). *Handleiding bij de herziene versie van Hare's Psychopathie Checklist*. Lisse: Swets Test Publishers.
- Vogel de, V., de Ruiter, C., Hildebrand, M., Bos, B., & van de Ven, P. (2004). Different ways of discharge and (risk of) recidivism measured by the HCR-20 and PCL-R in a sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health*, 3, 149-165.
- Vogel de, V. (2005). *Structured Risk Assessment of (Sexual) Violence in Forensic Clinical Practice. The HCR-20 and SVR-20 in Dutch Forensic Psychiatric Patients*. Amsterdam: Universiteit van Amsterdam [dissertatie].
- Wartna, B.S.J., el Harbachi, S., & van der Knaap, L.M. (2005). *Buiten behandeling: een cijfermatig overzicht van de strafrechtelijke recidive van ex-terbeschikkinggestelden*. WODC