Predictive Validity of the HKT-R Risk Assessment Tool: Two and 5-Year Violent Recidivism in a Nationwide Sample of Dutch Forensic Psychiatric Patients

Stefan Bogaerts[1,2], Marinus Spreen[3], Paul ter Horst[4], and Coby Gerlsma[5]

Abstract

This study has examined the predictive validity of the Historical Clinical Future [*Historisch Klinisch Toekomst*] Revised risk assessment scheme in a cohort of 347 forensic psychiatric patients, which were discharged between 2004 and 2008 from any of 12 highly secure forensic centers in the Netherlands. Predictive validity was measured 2 and 5 years after release. Official reconviction data obtained from the Dutch Ministry of Security and Justice were used as outcome measures. Violent reoffending within 2 and 5 years after discharge was assessed. With regard to violent reoffending, results indicated that the predictive validity of the Historical domain was modest for 2 (area under the curve [AUC] = .75) and 5 (AUC = .74) years. The predictive validity of the Clinical domain was marginal for 2 (admission: AUC = .62; discharge: AUC = .63) and 5 (admission: AUC = .69; discharge: AUC = .62) years after release. The predictive validity of the Future domain was modest (AUC = .71) for 2 years and low for 5 (AUC = .58) years. The total score of the instrument was modest for 2 years (AUC = .78) and marginal for 5 (AUC = .68) years. Finally, the Final Risk Judgment was modest for 2 years (AUC = .78) and marginal for 5 (AUC =

[1]Tilburg University, The Netherlands
[2]FPC Kijvelanden, Karid/Fivoor, Rotterdam, The Netherlands
[3]FPC Dr. S. van Mesdag Clinic, Groningen, The Netherlands
[4]FPK Woenselsepoort, Eindhoven, The Netherlands
[5]University of Groningen, The Netherlands

Corresponding Author:
Stefan Bogaerts, Department of Developmental Psychology, Tilburg University, Warandelaan 2, LE 5000 Tilburg, The Netherlands.
Email: s.bogaerts@uvt.nl

.63) years time at risk. It is concluded that this risk assessment instrument appears to be a satisfactory instrument for risk assessment.

Introduction

To assess the risk of violent reoffending in an individual patient after a forensic psychiatric treatment, forensic mental health professionals try to understand patient's typical behavioral mechanisms leading to future violent or aggressive behaviors (Bogaerts, Vanheule, & DeClercq, 2005; Bogaerts, Vervaeke, & Goethals, 2004; Buschman et al., 2010; Kunst, Bogaerts, & Winkel, 2009; Kunst, Winkel, & Bogaerts, 2010; Monahan, 1981). In risk assessment, typical behavioral mechanisms related to violent behavior express the vulnerability of the forensic patient to reoffend and are extrapolated and weighed for his likely future circumstances. In the Netherlands, risk assessment in forensic psychiatry has been mandatory since 2005. Patients convicted by the court to a TBS measure (*ter beschikkingstelling*: involuntary admission by order of the state; de Ruiter & Trestman, 2007) have committed a crime caused by their problematic mental state during a period of insanity. The primary goal of a TBS order is the protection of society and the resocialization of the patients into society (Brink, van den Hooijschuur, Os, van Savenije, & Wiersma, 2010; Van Marle, 2002). All forensic patients with a TBS order receive treatment to reduce the risk of future recidivism and to prepare them for reintegration into society (Van der Veeken, Lucieer, & Bogaerts, 2016). Forensic psychiatric centers differ fundamentally from prisons. In prisons, the focus is primarily on securing and retaliation with very few options for treatment. Otherwise, in forensic psychiatric centers, the emphasis is also on securing society but much more on the provision of treatment programs (Bonta, Harman, Hann, & Cormier, 1996; Dienst Justitiële Inrichtingen, 2014). In the Netherlands, five degrees of sanity can be distinguished: full responsible, slightly diminished responsible, diminished responsible, severely diminished responsible, and total absence of responsibility. All patients, regardless of their mental illness, can be held (partly) insane, which is different than in the United States, for example. Patients are usually first sentenced to prison before being transferred to a high security forensic psychiatric center (de Ruiter & Trestman, 2007; Radovic, Meynen, & Bennet, 2015).

To support and standardize individual risk assessment procedures, risk assessment instruments using Structured Professional Judgment (SPJ) procedures have become important to assess the likelihood of future violent recidivism (Douglas, Ogloff, Nicholls, & Grant, 1999; Hart, 1998; Webster, Douglas, Eaves, & Hart, 1997). SPJ instruments are composed of predetermined sets of empirically established static and dynamic risk factors associated with future reoffending and use specific and transparent guidelines. The rationale of a SPJ risk assessment instrument is that a higher total sum score is associated

with a higher likelihood of recidivism, but with the possibility for the clinician to over-rule the actuarial final assessment when there are good reasons for this. This positive linear association between the total sum score and the likelihood of recidivism must at least exist at the group level to ensure that each patient is evaluated with indicators, which are associated with recidivism (Dawes & Corrigan, 1974). However, in individual cases, this association is not necessarily positive linear. Therefore, in an individual risk assessment, assessors interpret and weigh a patient's scores not only on the established risk factors of the applied instrument, but they also consider specific indicators, risk management, and circumstances not captured by the instrument. The result of the combination and interpretation of all available information, preferably in consensus, is called the Final Risk Judgment (FRJ; Dawes, Faust, & Meehl, 1989; Spreen, Brand, Ter Horst, & Bogaerts, 2014). Examples of SPJ risk assessment schemes are the Historical Clinical Risk Management–20^{V3} (HCR-20; Douglas, 2014; Douglas, Hart, Webster, & Belfrage, 2013; Webster et al., 1997), the Short-Term Assessment of Risk and Treatability (START; Webster, Martin, Brink, Nicholls, & Middleton, 2004), and, in the Netherlands, the *Historisch Klinisch Toekomst*-30 (Historical Clinical Future–30 [HKT-30], and, its successor, the HKT-R; Spreen et al., 2014).

A risk assessment is not a single event within a patient's treatment trajectory but it has different functions dependent on the decision to make in treatment, such as whether a patient is ready, for example, to go with unguided leave for a short period. At the start of a treatment soon after the admission, the function of a risk assessment is to estimate future risk and to diagnose significant risk factors and tailor treatment goals to reduce these risk factors. Therefore, during treatment, risk assessments are important to evaluate whether there is a decrease of those risk factors. Throughout individual treatment trajectories, risk assessments can support decisions on, for example, leave or discharge modalities. In the Netherlands, it is imposed by the Ministry of Security and Justice that the risk of recidivism should be measured annually (Dienst Justitiële Inrichtingen, 2014; Spreen et al., 2014).

In risk assessment studies, the predictive validity is the main psychometric indicator to express the performance of a SPJ instrument in clinical practice (Singh, Grann, & Fazel, 2011). In the majority of risk assessment studies, a limited definition of predictive validity of measurements instruments is used, that is, the ability to discriminate between recidivists and non-recidivists. In this article, we apply the area under the curve (AUC) to evaluate the ability of the HKT-R to discriminate between recidivists and non-recidivists. The AUC value achieved from receiver operating characteristics (ROC) can range between 0 and 1, with .50 representing performance equivalent to change, and .75 and higher can be seen as a threshold for a large effect size (O'Shea & Dickens, 2015).

Predictive Validity of SPJ Instruments and Conditions

In the SPJ approach, a set of indicators are used that are theoretically and/or empirically related to future offending to indicate predictions of future recidivism (Singh, Serper, Reinharth, & Fazal, 2011). An important goal of risk assessments is to assess the risk of recidivism and to make a distinction between high-risk versus low-risk patients before

the start of the treatment and before release. This implies that an instrument must have discriminative power to make a distinction between recidivists and non-recidivists. Furthermore, risk assessment instrument must be sufficiently sensitive to measure behavioral and/or personal change to assess changes in risk levels during treatment (Jeandarme, Habets, Oei, & Bogaerts, 2017). These changes must be part of the risk assessment. For instance, suppose two patients have the same total score on a clinical scale of a risk assessment instrument at the end of their treatment, say 10. In mainstream predictive validity risk studies, both 10s are considered as equal in relation to relapse. However, if one patient had a score 30 and the other 15 at the start of their treatments, then the end score of 10 expresses different progress. Methodologically, this implies that to establish predictive validity of a risk assessment instrument, a patient's progress as measured by clinical scales must be weighted in the end total score expressed by the FRJ of the clinician to individualize the scores for each specific patient (Spreen et al., 2014).

Another aspect of predictive validity is the representativeness of the sample. The sample must be representative for the population and the population from which the predictive indicators are computed must resemble the current population as much as possible. Predictive values measured in a population in the eighties or nineties of the last century are not valid for a current population because society, treatments, and legislations have changed, and also scoring behavior can be changed over the years (Spreen et al., 2014). For example, in the Netherlands, the Ministry of Security and Justice mandated risk assessment in 2005, which means that each discharge of a forensic patient must be evidenced by a structured risk assessment, which is a completely different situation compared with the period before 2005.

This Study

The objective of this study is to investigate the predictive validity (discriminative power between recidivists and non-recidivists) of the Dutch risk assessment instrument HKT-R (Spreen et al., 2014). We will examine the predictive validity of the HKT-R for violent recidivism within 2 years and 5 years after release. Two measure points will be addressed, namely, the predictive validity at admission and discharge. First, the participants, procedure, outcome measures, instrument, and analysis methods will be discussed. Second, the predictive validity for violent reoffending within 2 and 5 years will be reported. We examine the predictive validity of the Historical, Clinical and Future domain; the HKT-R total score; and the FRJ during two measurement times, at the moment of admission and unconditional discharge. Because the HKT-R was developed to estimate violent recidivism, nonviolent, general, and sexual recidivism (child molestation) were excluded.

Materials and Methods

Background

The validation study of the HKT-R is the first nationwide study in the Netherlands that examined the predictive validity of the cohort forensic psychiatric TBS patients

discharged unconditionally between 2004 and 2008 (Spreen et al., 2014). The HKT-R is the recently revised version of the HKT-30 and developed in the Netherlands. The HKT-30 has been extensively studied in the past. For example, in a study of Canton et al. (2004), the predictive validity of the HKT-30 was examined in 123 cases. A distinction was made between serious and less severe recidivism. For serious recidivism, the predictive validity of the subscales was moderate (AUC Historical [H] scale = .69, AUC Clinical [K] scale = .68, and AUC Future [T] scale = .73). The predictive value of the subscales of the HKT-30 for less serious recidivism was low to moderate (AUC H-scale = .61, AUC K-scale = .59, and AUC T-scale = .68). The AUC values of the total score were not significantly different from serious and less severe recidivism (Canton et al., 2004).

Participants

The original study sample ($N = 347$) consisted of all Dutch forensic psychiatric patients who were discharged between 2004 and 2008 from any of 12 Dutch forensic psychiatric hospitals. Fifteen male patients (4.3%) were not involved in the study: possible reconviction data of five patients (1.4%) could not be found, and 10 patients (2.9%) died within 2 years after discharge. Therefore, the final sample consisted of 332 patients. Assessment refers to the period of admission and unconditional release.

Procedure

The HKT-R was retrospectively coded using information derived from the patients' criminal files. In general, these criminal files were extensive and contained detailed descriptions of the patients' background and criminal history, psychiatric evaluation reports, treatment plans, leave requests, and prolongation advices. Raters consisted of 20 graduate students in Forensic Psychology who had received an intensive, 2-week training in advance. Sixty random files ($N = 12$, random from each of the five discharge years 2004 to 2008) were doubly scored by two independent raters to establish the interrater reliability, which gives an indication of the degree to which two raters assign the same scores for the assessment of the same patient. The interrater reliability of the various components of the HKT-R was assessed on the basis of the one-way random, single measure intraclass correlation coefficient (ICC). ICC values of <.40 are generally considered low, ICC values from .40 to .74 as reasonable to good, and values of ≥.75 as very good (Fleiss, 1986). The interrater reliability for the Historical domain was .80 (.66-.89), for the Clinical domain .85 (.67-.94), and for the Future domain .42 (.09-.67). The interrater reliability for the total instrument was .62 (.41-.77).

Outcome Measures of Reoffending

Official reconviction data were obtained from the Dutch Ministry of Security and Justice. In the database that was provided, the patients had been tracked from discharge until July 11, 2011. Based on this information, the outcome measures "violent

reoffending within two and five years" were constructed. Violent reoffending was operationalized as any new conviction for an offense involving moderate violence, robbery with violence, serious violence, and arson with the risk for life, (attempted) homicide/murder, and violent sexual assaults on adult victims.

Instrument

The HKT-R is a structured professional risk assessment instrument that assesses 12 Historical, 14 Clinical, and seven Future risk factors for violent reoffending in forensic psychiatric patients. All those risk factors are associated with future recidivism (Andrews & Bonta, 2006; Andrews, Bonta, & Wormith, 2011). The Historical domain regards the past of the patient, up to the time of the arrest. The Clinical domain is related to the behavior of the patient during (the last 12 months) forensic psychiatric treatment. The Future domain refers to the situation the patient returns to after discharge from the forensic psychiatric hospital and expectations with respect to his ability to retain his newly acquired skills (Spreen et al., 2014). All individual items are measured on a 5-point scale ranging from 0 to 4, with higher scores indicating that a particular risk factor is more present in the patient being assessed. The scores on the individual items are summed up into domain scores and a HKT-R total score (in theory ranging from 0 to 132). Finally and importantly, the professional judgment is used to weight and interpret the actuarial information into an FRJ by the clinicians, expressed in terms of low, low/medium, medium, medium/high, and high risk.

Statistical Analyses

To discriminate between violent recidivists and nonviolent recidivists, we expressed the degree of predictive validity by using ROCs analyses, resulting in AUC values. An AUC value of .50 implies prediction equal to chance whereas an AUC value of 1.00 represents perfect prediction. Intermediary AUC values were interpreted as follows: .51-.60 low accuracy; .60-.70 marginal accuracy; .70-.80 modest accuracy; .80-.90 moderate accuracy; and over .90 high accuracy (Sjöstedt & Grann, 2002). In this study, we examined 2- and 5-year violent recidivism.

Results

Group Description

The majority of patients who were included in the current study were male ($n = 305$, 91.9% vs. $n = 27$, 8.1% female). Most patients were born in the Netherlands ($n = 254$, 76.5%), whereas the other patients were born in Suriname ($n = 30$, 9.0%), the Netherlands Antilles ($n = 18$, 5.4%), or elsewhere ($n = 30$, 9.0%). The index offenses for which the mandatory treatment order had been imposed included manslaughter ($n = 105$, 31.6%), murder ($n = 57$, 17.2%), sex offenses ($n = 47$, 14.2%), arson ($n = 40$, 12.0%), robbery ($n = 40$, 12.0%), severe violence ($n = 24$, 7.2%), and other violent

offenses ($n = 19$, 5.7%). The mean age at admission to the forensic psychiatric hospital was 31.84 years ($SD = 8.58$, 16.32-64.81). The most prevalent *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association, 1994) diagnoses, which had been assessed at the start of treatment, were personality disorder Not Otherwise Specified (NOS; $n = 171$, 51.5%), substance-related disorders ($n = 111$, 33.4%), Cluster B personality disorders ($n = 93$, 28.0%), and schizophrenia and other psychotic disorders ($n = 68$, 20.5%). These percentages do not count to 100% because in most patients, comorbid disorders were present. The mean total IQ of the study sample was 98.32 ($SD = 15.72$, 52-139).

After a mean period of residence of 8.21 years ($SD = 3.24$, 2.63-27.68), the mean age at discharge from the psychiatric hospital was 40.03 years ($SD = 8.88$, 24.76-69.43). After discharge, the majority of patients ($n = 236$, 71.1%) returned to live independently in the community, whereas the remainder of the patients returned to supervised settings, such as supervised living facilities ($n = 72$, 21.7%), general psychiatric hospitals ($n = 15$, 4.5%), institutions for the mentally disabled ($n = 8$, 2.4%), and nursing homes ($n = 1$, 0.3%).

Prevalence of Reoffending

After a mean follow-up of 2 years, 41 (12.3%) patients of 332 patients who could be located after 2 years had recidivated at least one time with a violent offense; 36 patients (19.3%) of the 187 patients who could be traced after 5 years time at risk had recidivated at least once with a violent offense. After 2 years, 15 male patients of the 347 patients were missing although all 347 patients had a time at risk of 2 years. The Dutch Ministry of Security and Justice could not precisely specify the reasons for the missing persons. It was assumed that some were deceased during the time at risk, some might be facing abroad, others were detained, and others had no official residential address. None of the women in the cohort had relapsed.

Predictive Validity of the HKT-R

The predictive validity of the different domains of the HKT-R for violent reoffending will be investigated for a 2- and 5-year period time at risk.

Violent reoffending within 2 years. The Historical domain (AUC = .75%, confidence interval [CI] = [.64, .87]) was found to be modestly predictive of violent reoffending after a period of 2-year release; the Clinical domain at admission (AUC = .62%, CI = [.48, .77]) and at discharge (AUC = .63%, CI = [.48, .77]) were marginally predictive of violent recidivism. The Future domain (AUC = .71%, CI = [.59, .84]), the HKT-R total score (AUC = .78%, CI = [.67, .89]), and the HKT-R FRJ (AUC = .78%, CI = [.67, .89]) were found to be modest predictors of violent reoffending after 2 years of discharge. Scores on the respective scales of the HKT-R, the total scores on the HKT-R, and the HKT-R FRJ score were all significantly different between recidivists and non-recidivists: (H-total, $\mu_{recidivists} = 30.48$, $SD = 8.42$; $\mu_{non\text{-}recidivists} = 21.60$, $SD = 8.52$,

$p < .001$; K-total admission, $\mu_{recidivists} = 22.43$, $SD = 10.49$; $\mu_{non-recidivists} = 18.43$, $SD = 8.43$, $p < .05$; K-total discharge, $\mu_{recidivists} = 12.42$, $SD = 10.72$; $\mu_{non-recidivists} = 8.09$, $SD = 6.85$, $p < .05$; T-total, $\mu_{recidivists} = 8.37$, $SD = 6.29$; $\mu_{non-recidivists} = 6.06$, $SD = 4.88$, $p < .05$; HKT-total, $\mu_{recidivists} = 51.46$, $SD = 19.49$; $\mu_{non-recidivists} = 35.19$, $SD = 15.33$, $p < .001$; and HKT-R FRJ, $\mu_{recidivists} = 2.90$, $SD = 1.36$; $\mu_{non-recidivists} = 2.07$, $SD = 1.15$, $p < .001$).

Violent reoffending within 5 years. After 5 years time at risk, the Historical domain was found to be a modestly accurate predictor of violent recidivism (AUC = .74%, CI = [.55, .94]). The Clinical domain at admission (AUC = .69%, CI = [.53, .85]) and at discharge (AUC = .62%, CI = [.44, .80]) marginally predicted recidivism after a time at risk of 5 years. The Future domain was found to be low predictive (AUC = .58%, CI = [.38, .79]). The total HKT-R score was a marginal predictor of violent reoffending after 5 years of unconditional release (AUC = .68%, CI = [.47, .90]), and the HKT-R FRJ was also marginal (AUC = .63%, CI = [.41, .84]). Scores on the respective scales of the HKT-R, the total scores on the HKT-R, and the HKT-R FRJ score were all significantly different between recidivists and non-recidivists: (H-total, $\mu_{recidivists} = 28.86$, SD = 7.29; $\mu_{non-recidivists} = 20.48$, SD = 8.10, p < .001; K-total admission, $\mu_{recidivists} = 22.63$, SD = 10.55; $\mu_{non-recidivists} = 18.38$, SD = 8.30, p < .05; K-total discharge, $\mu_{recidivists} = 11.11$, SD = 8.83; $\mu_{non-recidivists} = 7.37$, SD = 6.88, p < .05; T-total, $\mu_{recidivists} = 7.29$, SD = 5.21; $\mu_{non-recidivists} = 5.45$, SD = 4.59, p < .05; HKT-total, $\mu_{recidivists} = 48.14$, SD = 16.21; $\mu_{non-recidivists} = 31.90$, SD = 13.70, p < .001; and HKT-R FRJ, $\mu_{recidivists} = 2.72$, SD = 1.30; $\mu_{non-recidivists} = 2.09$, SD = 1.18, p < .001).

In Table 1, the correlations were shown between the Historical, Clinical, Future, HKT-R total, and HKT-R FRJ with violent recidivism after 2 and 5 years time at risk. All correlations were significant. Based on the results from the ROC analysis, AUC values for the Clinical and Future domain and HKT-R-total were similar or even better for 5-year than 2-year recidivism. Only the Historical domain and the FRJ showed a lower AUC value after 5 years compared with 2-year recidivism.

Discussion

In this study, we examined the predictive validity of the HKT-R risk assessment scheme for violent reoffending in a sample of 332 forensic psychiatric patients in the Netherlands who were unconditionally released between 2004 and 2008. The follow-up period was 2 and 5 years. The results in the study indicated that the predictive validity of the Historical domain of the HKT-R was modest for violent recidivism after 2 and 5 years time at risk. The predictive validity of the Clinical domain after 2 and 5 years time at risk was marginal and the predictive domain of the Future domain was also marginal. The HKT-R total score was modest after 2 and 5 years time at risk; the FRJ after 2 years was modest and after 5 years marginal.

Comparing these results with those found in other studies using the HKT-30 (Canton et al., 2004; Hildebrand, Hesper, Spreen, & Nijman, 2005; Schönberger, Hildebrand, Spreen, & Bloem, 2008; van den Brink et al., 2010), the predictive

Table 1. The Predictive Validity of the HKT-R for 2- and 5-Year Violent Recidivism and Point Biserial Correlation Coefficients.

Domain (label)	Violent reoffending ≤2 years				Violent reoffending ≤5 years			
	AUC	SE	95% CI	R_{pb}	AUC	SE	95% CI	R_{pb}
Historical	.75	.06	[.64, .87]	.30**	.74	.10	[.55, .94]	.11
Clinical admission	.62	.07	[.48, .77]	.17*	.69	.08	[.53, .85]	.20*
Clinical discharge	.63	.08	[.48, .77]	.26**	.62	.09	[.44, .80]	.40**
Future	.71	.06	[.59, .84]	.28**	.58	.10	[.38, .79]	.36**
HKT-R total score	.78	.06	[.67, .89]	.36**	.68	.11	[.47, .90]	.36**
Final Risk Judgment	.78	.05	[.67, .89]	.35**	.63	.11	[.41, .84]	.27**

Note. HKT-R = *Historisch Klinisch Toekomst*; AUC = are under the curve; CI = confidence interval; R = Pearson's correlation.
*p < .05. **p < .01 (two-tailed; point biserial correlation coefficient).

validity of the Historical domain of the HKT-R has clearly improved; the predictive validity of the Clinical and Future domains have stayed stable. The predictive accuracy of the total score stayed stable and the FRJ after 5 years has somewhat decreased. However, other studies were all based on selective and smaller samples what makes it difficult to compare our results . Furthermore, results found in other studies differ on study designs, sample sizes, patient characteristics, outcome measures of reoffending, and follow-up term.

Strength and Limitations of the Study

Several issues must be mentioned. First, this study was based on a saturated nationwide cohort of forensic psychiatric patients discharged between 2004 and 2008 from one of the 12 TBS hospitals in the Netherlands. It is the first study in the Netherlands that is not based on a limited and selected non-representative sample. This study is in accordance with the requirements asked by the Dutch Institute of Psychologists, namely, that only nationally validated instruments may be used in clinical research of (forensic) patients. A second remark concerns the retrospective study design in which the HKT-R was assessed during admission and discharge using official criminal file information. Although these criminal files were, in general, extensive and contained a lot of relevant information, it may have been better to assess the HKT-R on direct observations of the patient. In addition, it is also preferred to report measurements when patients, for example, deal with unguided leave. A third remark concerns the outcome, namely, recidivism. It is important to not only look at official reconvictions but also take into account arrests. A last concern relates the evidence of the HKT-R's calibration, which means the ability of a risk assessment tool to estimate rates of recidivism for single risk scores. In this study, we have not paid attention to the calibration as mentioned by Rossegger, Endrass, Gerth, and Singh (2014).

Conclusion

The HKT-R appears to be a good revised instrument for the assessment of future violent risk, although more studies are needed to replicate these findings in different samples. In future studies, it would be important to examine the predictive validity of the HKT-R in specific groups of patients, for instance, on the basis of their gender, psychopathology, offense history, and violent behavior during treatment. In addition, it would be valuable to assess changes in dynamic indicators of the HKT-R over time to examine how these changes relate to post-discharge reoffending (Schuringa, Spreen, & Bogaerts, 2014). In the end, these efforts may not only result in more accurate predictions of the risk of recidivism in forensic psychiatric patients but also give treatment indications of which risk factors should be treated and which protective factors should be continued and strengthened. Finally and related to clinical implications, risk assessment is a complex task and must be in function of understanding risk at an individual level to develop management plans. The HKT-R can be seen as a useful and validated instrument to assist in this process.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Cincinnati, OH: Anderson.

Andrews, D. A., Bonta, J., & Wormith, S. J. (2011). The risk-need-responsivity (RNR) model. *Criminal Justice and Behavior*, *38*, 735-755.

Bogaerts, S., Vanheule, S., & DeClercq, F. (2005). Recalled parental bonding, adult attachment style, and personality disorders in child molesters: A comparative study. *The Journal of Forensic Psychiatry & Psychology*, *16*, 445-458.

Bogaerts, S., Vervaeke, G., & Goethals, J. (2004). A comparison of relational attitude and personality disorders in the explanation of child molestation. *Sexual Abuse: A Journal of Research and Treatment, 16*, 37-47.

Bonta, J., Harman, W. G., Hann, R. G., & Cormier, R. B. (1996). The prediction of recidivism among federally sentenced offenders: A revalidation of the SIR scale. *Canadian Journal of Criminology*, *38*, 61-79.

Buschman, J., Bogaerts, S., Foulger, S., Wilcox, D., Sosnowski, D., & Cushman, B. (2010). Sexual history disclosure polygraph examinations with cybercrime offences: A first Dutch explorative study. *International Journal of Offender Therapy and Comparative Criminology*, *54*, 395-411.

Canton, W. J., Veer, T. S., van der Panhuis, P. J. A., van Verheul, R., & van den Brink, W. (2004). De voorspellende waarde van risicotaxatie bij de rapportage *pro Justitia*: Onderzoek naar de HKT-30 en de klinische inschatting [The predictive validity of risk assessment in "*pro Justitia*" reports. A study of the HKT-30 and clinical assessment]. *Tijdschrift voor Psychiatrie*, *46*, 525-535.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95-106.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668-1674.

de Ruiter, C., & Trestman, R. L. (2007). Prevalence and treatment of personality disorders in Dutch forensic mental health services. *Journal of the American Academy of Psychiatry Law*, *35*, 92-97.

Dienst Justitiële Inrichtingen. (2014). *Kernset prestatie-indicatoren forensische psychiatrie verslagjaar 2014* [Core set of performance indicators for forensic psychiatry]. Retrieved from http://www.forensischezorg.nl/files/indicatorengids_verslagjaar_2014_v1.1.pdf

Douglas, K. S. (2014). Version 3 of the Historical-Clinical-Risk Management-20 (HCR-20(V3)): Relevance to violence risk assessment and management in forensic conditional release contexts. *Behavioral Sciences & the Law*, *32*, 557-576.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3 assessing risk for violence.* Vancouver, British Columbia, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., Ogloff, R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 violence risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology*, *67*, 917-930.

Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York, NY: John Wiley.

Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology*, *3*, 121-137.

Hildebrand, M., Hesper, B. L., Spreen, M., & Nijman, H. L. I. (2005). *De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie* [The value of structured professional risk assessment and the diagnosis psychopathy]. Utrecht, The Netherlands: Expertisecentrum Forensische Psychiatrie.

Jeandarme, I., Habets, P., Oei, T. I., & Bogaerts, S. (2017). PCL-R field validity in prison and hospital settings. *Law and Human Behavior*, *41*, 29-43.

Kunst, M. J. J., Bogaerts, S., & Winkel, F. (2009). Peer and inmate aggression, type D-personality and posttraumatic stress among Dutch prison workers. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *25*, 387-395.

Kunst, M. J. J., Winkel, F. W., & Bogaerts, S. (2010). Posttraumatic growth moderates the association between violent revictimisation and persisting PTSD symptoms in victims of interpersonal violence: A six-month follow-up study. *Journal of Social & Clinical Psychology*, *29*, 527-545.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverley Hills, CA: SAGE.

O'Shea, L. E., & Dickens, G. L. (2015). Predictive Validity of the Short-Term Assessment of Risk and Treatability (START) for aggression and self-harm in a secure mental health service: Gender differences. *International Journal of Forensic Mental Health*, *14*, 132-146.

Radovic, S., Meynen, G., & Bennet, T. (2015). Introducing a standard of legal insanity: The case of Sweden compared to The Netherlands. *International Journal of Law and Psychiatry*, *40*, 43-49.

Rossegger, A., Endrass, J., Gerth, J., & Singh, J. P. (2014). Replicating the violence risk appraisal guide: A total forensic cohort study. *PLoS ONE*, *9*(3), e91845.

Schönberger, H. J. M., Hildebrand, M., Spreen, M., & Bloem, O. (2008). *De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie bij seksuele delinquenten* [The value of structured professional risk assessment and the diagnosis psychopathy in sex offenders]. Utrecht, The Netherlands: Expertisecentrum Forensische Psychiatrie.

Schuringa, E., Spreen, M., & Bogaerts, S. (2014). Inter-rater and test-retest reliability, internal consistency and factorial structure of the Dutch Instrument for Forensic Treatment Evaluation. *Journal of Forensic Psychology Practice*, *14*, 127-144.

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

Singh, J. P., Serper, M., Reinharth, J., & Fazal, S. (2011). Structured assessment of violence risk in schizophrenia and other psychiatric disorders: A systematic review of the validity, reliability, and item content of 10 available instruments. *Schizophrenia Bulletins*, *37*, 899-912.

Sjöstedt, G., & Grann, M. (2002). Risk assessment: What is being predicted by actuarial prediction instruments? *International Journal of Forensic Mental Health*, *1*, 179-184.

Spreen, M., Brand, E., Ter Horst, P., & Bogaerts, S. (2014). *Handleiding en Methodologische Verantwoording HKT-R, Historisch, Klinische en Toekomstige – Revisie* [Guidlines and Methodological Research of the HKT-R, Historical, Clinical and Future – Revision]. Groningen, The Netherlands: Dr. van Mesdag kliniek.

van den Brink, R. H. S., Hooijschuur, A., Os, T. W. D. P., van Savenije, W., & Wiersma, D. (2010). Routine violence risk assessment in community forensic mental healthcare. *Behavioral Sciences & the Law*, *28*, 396-410.

Van der Veeken, F., Lucieer, J., & Bogaerts, S. (2016). Routine outcome monitoring and clinical decision-making in Forensic Psychiatry based on the Instrument for Forensic Treatment Evaluation. *PLoS ONE*, *11*, e0160787.

Van Marle, H. J. C. (2002). The Dutch Entrustment Act (TBS): Its principles and innovations. *International Journal of Forensic Mental Health*, *1*, 82-92.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. (1997). *HCR-20: Assessing risk for violence, Version 2*. Burnaby, British Columbia, Canada: Simon Fraser University.

Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Middleton, C. (2004). *Manual for the Short-Term Assessment of Risk and Treatability (START)*. Hamilton, Ontario, Canada: St. Joseph's Healthcare, Hamilton and British Columbia Mental Health and Addiction Services.