

**HCR-20 VIOLENCE RISK ASSESSMENT SCHEME:
OVERVIEW AND ANNOTATED BIBLIOGRAPHY
(WITH SUPPLEMENT, CURRENT UP TO JANUARY 12, 2006)**

**KEVIN S. DOUGLAS, LL.B., PH.D.
LAURA S. GUY, M.A.
SIMON FRASER UNIVERSITY**

**JOHN WEIR, M.S.
UNIVERSITY OF SOUTH FLORIDA**

(Visit <http://www.sfu.ca/psychology/groups/faculty/hart/violink.htm> for updates)

Check out the Annotated Bibliography for the Sexual Violence Risk – 20 (SVR-20), at the above web address!
SVR-20 Annotated Bibliography Prepared by Vivienne de Vogel, Ph.D. (vdevogel@hoevenstichting.nl)

*****NEW ADDITIONS*****

1. Belfrage, H. Fransson, G., & Strand, S. (2004). Management of violent behavior in the correctional system using qualified risk assessments. *Legal and Criminological Psychology*, 9, 11-22. (p. 38)
2. Brown, L. K. (2004). *Assessing risk for elopement and breaches of conditional release in insanity acquittees*. Unpublished Ph.D. Dissertation. Simon Fraser University, Burnaby, British Columbia, Canada. (p. 16)
3. de Vogel, V. & de Ruiter, C. (in press). The HCR-20 in personality disordered female offenders: A comparison with a matched sample of males. *Clinical Psychology and Psychotherapy*. (p. 23)
4. de Vogel, V., & de Ruiter, C. (under review). *Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20*. Manuscript submitted for publication. (p. 24)
5. Dolan, M., & Khawaja, A. (2004). The HCR-20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior*, 30, 469-483. (p. 19)
6. Douglas, K. S., & Ogloff, J. R. P. (2003). The impact of confidence on the accuracy of structured professional and actuarial violence risk judgments in a sample of forensic psychiatric patients. *Law and Human Behavior*, 27, 573-587. (p. 21)
7. Folino, J. O., Marengo, C. M., Marchiano, & Ascazibar, M. (2004). The risk assessment program and the Court of Penal Execution in the province of Buenos Aires, Argentina. *International Journal of Offender Therapy and Comparative Criminology*, 48, 49-58. (p. 46)
8. Fujii, D., Lichten, A., & Tokioka, A. (under review). Structured professional judgment versus actuarial data in violence risk prediction using the Historical Clinical Risk Management-20. Manuscript under review. (p. 27)

9. Fujii, D., Tokioka, A., & Lichten, A. (in press). Cultural differences in violence risk prediction. (in press). Cultural differences in violence risk prediction of psychiatric inpatients using the Historical Clinical Risk Management-20. *Psychiatric Services*. (p. 27)
10. Gray, N. S., Hill, C., McGleish, A., Timmons, D., MacCulloch, M. J., & Snowden, R. J. (2003). Prediction of violence and self-harm in mentally disordered offenders: A prospective study of the efficacy of the HCR-20, PCL-R, and psychiatric symptomatology. *Journal of Consulting & Clinical Psychology*, *71*, 443-451. (p. 42)
11. Gray, N. S., Snowden, R. J., MacCulloch, S., Phillips, H., Taylor, J. & MacCulloch, M. J. (2004). Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally disordered offenders: A comparative study of HCR-20, PCL:SV, and OGRS. *Journal of Consulting and Clinical Psychology*, *72*, 523-530. (p. 29)
12. Grevatt, M., Thomas-Peter, B., & Hughes, G. (2004). Violence, mental disorder and risk assessment: Can structured clinical assessments predict the short-term risk of inpatient violence? *The Journal of Forensic Psychiatry & Psychology*, *15*, 278-292. (p. 30)
13. Urheim, R., Jakobsen, D., & Rasmussen, K. (2003, August). *Dimensions of inpatient aggressive behavior in a security ward: What is being "predicted"?* Paper presented at the 5th Nordic Symposium on Forensic Psychiatry, Ystad, Sweden. (p. 35)
14. Watt, A., Topping-Morris, B., Rogers, P., Doyle, M., & Mason, T. (2003). Pre-admission nursing assessment in a Welsh medium secure unit (1991-2000): Part 2 – comparison of traditional nursing assessment with the HCR-20 risk assessment tool. *International Journal of Nursing Studies*, *40*, 657-662. (p. 36)

DESCRIPTION OF THE HCR-20

VIOLENCE RISK ASSESSMENT SCHEME

The HCR-20 (Webster, Douglas, Eaves, & Hart, 1997a; see Webster, Eaves, Douglas, & Wintrup, 1995, for Version 1) is a broad-band violence risk assessment instrument with potential applicability to a variety of settings. The conceptual scheme of the HCR-20 aligns risk markers into past, present, and future. Its 10 Historical factors obviously concern the past. However, the HCR-20 contains 5 Clinical items that are meant to reflect current, dynamic (changeable) correlates of violence. The future is recognized in the 5 Risk Management items, which focus attention on situational post-assessment factors that may aggravate or mitigate risk. The HCR-20 takes its name from these three scales — Historical, Clinical, Risk Management — and from the number of items (20). Table 1 shows the items.

The HCR-20 was developed from a thorough consideration of the empirical literature concerning factors that relate to violence. It attempts to develop professional standards regarding the process and substance of risk assessments. Further, the HCR-20 integrates the experience of clinicians, and is easy to administer, understand, and score. Randy Borum (1996) recently has written about the HCR-20 that “the promise of this instrument lies in its foundation on a conceptual model or scheme for assessing dangerousness and risk; its basis in the empirical literature; its operationally defined coding system...[and] its practical use....The field eagerly awaits new data on this instrument” (p. 950).

Very complicated schemes may not be put to their intended use in the daily practice of risk assessment. Professionals who make risk assessments cannot afford the time to calculate complex weighting co-efficients

and discriminant function equations. As such, the HCR-20 is an attempt to merge science and practice by offering an instrument that can be integrated into clinical practice but also is empirically based and testable.

Table 1
Items in the HCR-20 Risk Assessment Scheme

Sub-Scales	Items
Historical Scale	
H1	Previous Violence
H2	Young Age at First Violent Incident
H3	Relationship Instability
H4	Employment Problems
H5	Substance Use Problems
H6	Major Mental Illness
H7	Psychopathy
H8	Early Maladjustment
H9	Personality Disorder
H10	Prior Supervision Failure
Clinical Scale	
C1	Lack of Insight
C2	Negative Attitudes
C3	Active Symptoms of Major Mental Illness
C4	Impulsivity
C5	Unresponsive to Treatment
Risk Management Scale	
R1	Plans Lack Feasibility
R2	Exposure to Destabilizers
R3	Lack of Personal Support
R4	Noncompliance with Remediation Attempts
R5	Stress

Note. Adapted from Webster, Douglas, Eaves, and Hart (1997a).

RECEIVER OPERATING CHARACTERISTIC (ROC) ANALYSES: AN EMERGING TECHNIQUE IN RISK ASSESSMENT RESEARCH

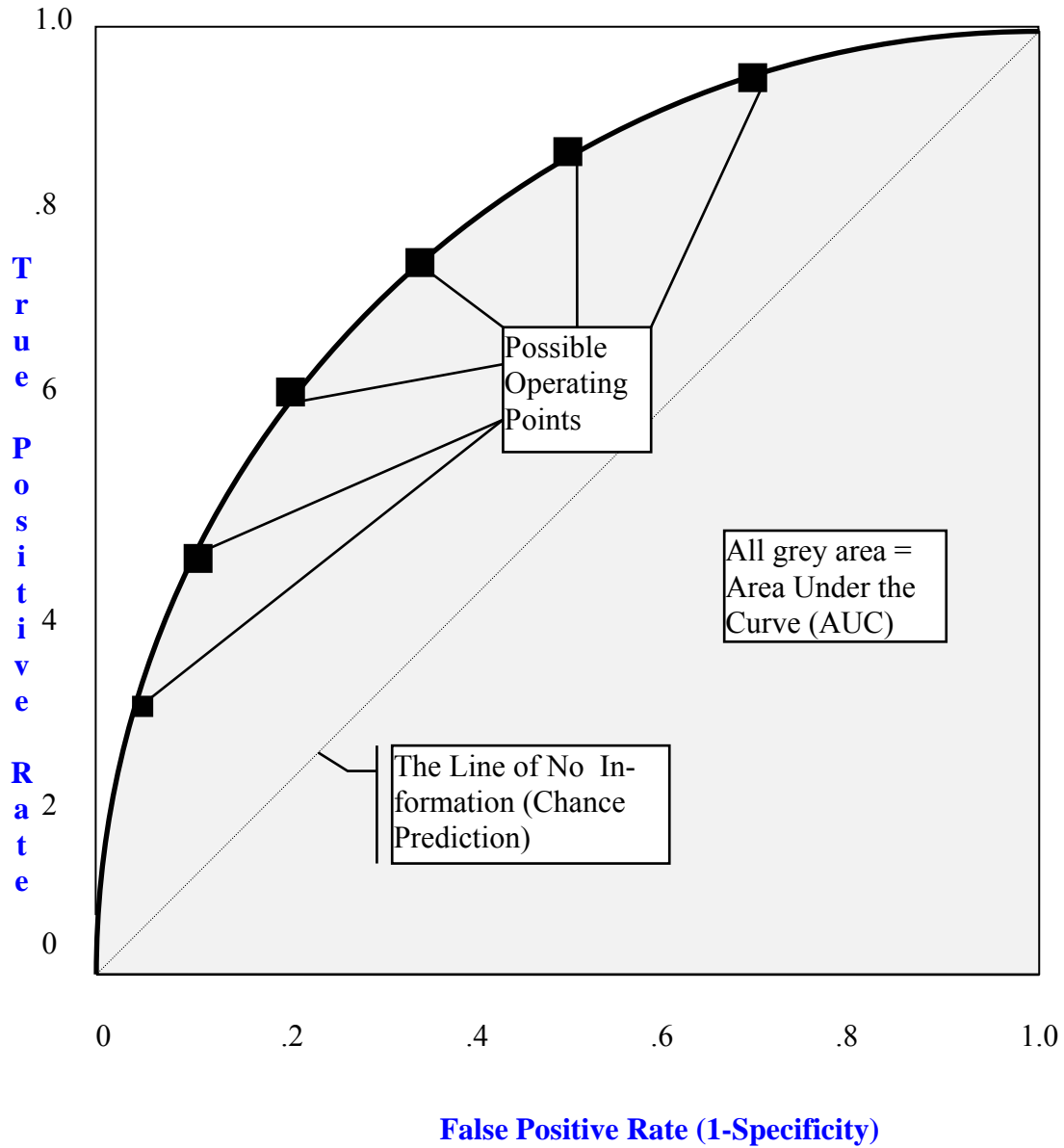
ROC statistical analysis is summarized here because several of the HCR-20 studies use this analysis, and results are reported in terms of the statistical indexes that ROC produces. Although ROCs have been used in the area of radiology (Lusted, 1978), radar signal detection, and sensory psychology since the 1950s and 1960s (Metz, 1984), they have only recently been introduced into the area of violence risk assessment (Mossman, 1994a, 1994b; Rice & Harris, 1995; Rice, 1997). They are recommended in this area because they are less dependent on the base rate of the criterion variable in the sample (in the present case, violence) than are traditional measures of predictive accuracy derived from 2 x 2 contingency tables (such as false positives and false negatives). Since correlations diminish with departures from base rates of 50%, correlational techniques are not the most effective means to estimate predictive efficiency of risk assessment schemes (Rice & Harris, 1995).

ROCs allow for the comparison of various thresholds on the predictor measures for offering predictions of violence, an overall index of accuracy which accounts for all possible thresholds, the simple identification of the optimal threshold, and the comparison of two or more predictors (Hsiao, Bartko, & Potter, 1989; Lusted, 1978; Metz, 1984; Mossman, 1994a; 1994b; Mossman & Somoza, 1991; Vida, 1997).

The term “receiver operating characteristic” took its name because it describes the detection, or prediction, “characteristics” of the test, and the “receiver” of the data can “operate” at any given point on the curve (Metz, 1978). ROCs are meant to be applied to data that are comprised of a continuous predictor variable and a dichotomous dependent measure. They take the form of a figure (see Sample ROC, next page, for an example) with the sensitivity (true positive rate [TPR]) of the predictor plotted as a function of the false positive rate (FPR [1-specificity]) (Mossman & Somoza, 1991). For any given level of specificity, the receiver knows the sensitivity. Each point on the curve (which corresponds to a cut-off on the predictor) represents a different trade-off between sensitivity and specificity.

The area under the curve (AUC) of the ROC graph can be taken as an index for interpreting the overall accuracy of the predictor. Areas can range from 0 (perfect negative prediction), to .50 (chance prediction), to 1.0 (perfect positive prediction). A given area represents the probability that a randomly chosen person who scores positive on the dependent measure (in this study, is actually violent) will fall above any given cut-off on the predictor measure, and that an actually non-violent person will score below the cut-off (Mossman & Somoza, 1991). Thus, an area of .75 means that there is a 75% chance that an actually violent person would score above the cut-off for violence on the predictor, and an actually non-violent person would score below the cut-off. AUC values of 0.70 may be considered moderate to large, and .75 and above may be considered large.

Figure 1
A Sample ROC Curve



A SUMMARY OF HCR-20 RESEARCH

STUDY/ SAMPLE	N	MEANS (SD)				VALIDITY INDICES	RELIABILITY INDICES
		HCR TOTAL	H SCALE	C SCALE	R SCALE		
CIVIL PSYCHIATRIC							
DOUGLAS, OGLOFF, NICHOLLS, GRANT (1999) ¹	193	19.0 (5.8)	9.8 (3.4)	4.1 (1.9)	5.1 (2.3)	AUCs = .76 - .80	IRR (TOTAL) = .80 (ICC)
MCNIEL ET AL. (IN PRESS)	100	18.0 (6.6)	7.1 (3.5)	6.1 (2.3)	4.8 (2.3)	AUC (TOT) = .65; (H,C,R,) = .56, .77, .58	ICC (TOTAL) = .78-.96
NICHOLLS, OGLOFF, DOUG- LAS (2001) ²	279	20.4 (5.6) ^A 16.8 (5.4) ^B	10.8 (3.3) ^A 8.2 (3.2) ^B	7.4 (1.5) ^A 6.9 (1.7) ^B	5.4 (2.4) ^A 4.8 (2.1) ^B	AUCs = .72 - .75 ^A AUCs = .66 - .80 ^B	SEE DOUGLAS ET AL (1999)
ROSS, HART, WEBSTER (1998) ³	131	19.1 (6.2)	8.7 (4.0)	5.6 (1.9)	4.8 (2.1)	AUCs (TOT) = .68 - .75	IRR (H) = .82; α (H) = .74; (C) = .64
FORENSIC PSYCHIATRIC							
BELFRAGE (1998)	43	28.8 (6.2)	13.8 (3.3)	5.5 (2.2)	6.6 (2.0)	NA	IRR (TOTAL) = .81 α (TOTAL) = .95
BROWN (2001)	172	22.7 (6.5)	12.8 (3.6)	4.5 (2.5)	5.4 (2.7)		ICC(H SCALE) = .80
CLAIX ET AL (2002)	86	23.3 (6.3)	12.4 (3.8)	5.0 (2.2)	6.0 (2.0)	r_{TOT} = .30 W/ ASSUALT	IRR (TOTAL) = .73
DERNEVIK (1998)	6*6	--	--	--	--	NA	IRR = .76 - .96
DERNEVIK ET AL (2001)	8 ¹¹	22.7 (6.5) ¹² 26.3 (6.1) ¹²	N/A	N/A	N/A	MULT R = .66 B/W HCR & FEELING CHECKLIST	N/A
DERNEVIK ET AL. (2002)	54	23.1 (5.8)	12.6 (3.5)	5.7 (2.1)	4.8 (1.7)	r_{TOT} = .32 (INPATIENT) AUC = .84 (COMMUNITY)	
DE VOGEL ET AL (2001)	60	26.1 (6.5)	14.6 (3.3)	5.3 (2.2)	6.1 (2.1)	NA	IRR (TOTAL) = .79

CONTINUES ON NEXT PAGE WITH MORE FORENSIC SAMPLES

STUDY/ SAMPLE	N	MEANS (SD)				VALIDITY INDICES	RELIABILITY INDICES
		HCR TOTAL	H SCALE	C SCALE	R SCALE		
FORENSIC PSYCHIATRIC							
DE VOGEL ET AL (2004)	120	22.8 – 32.0 ¹⁹	12.6 – 16.0	3.7 – 7.0	5.6 – 9.1	AUC (TOTAL) = .82 AUC (SPJ) = .79	IRR (TOTAL) = .83 IRR (SPJ) ²⁰ = .73
DOLAN & KHAWAJA (2004)	70	19.4 (5.7) ²⁴	11.8 (3.7)	3.3 (2.2)	4.1 (1.5)	AUC (TOT) = .67-.85 ²⁵	
DOUGLAS ET AL. (1998)	175	24.6 (5.8)	12.5 (3.6)	5.5 (2.5)	6.6 (2.3)	ODDS = 2.2 – 3.7	α (TOTAL) = .78 IRR (ICC) = .81
DOUGLAS ET AL. (2003) ¹⁸	100	24.7 (4.6)	14.4 (2.8)	4.7 (2.0)	5.9 (1.5)	AUC (TOT) = .67-.70 AUC (SPJ) = .68-.74	IRR (TOTAL) = .85 IRR (SPJ) = .61
FUJII, LICHTON ET AL. (2004)	169	--	--	--	--	AUC (TOT) = .61 ²⁶ AUC (SPJ) = .70	IRR (12 CASES) = .94
FUJII, TOKIOKA ET AL. (2004)	169	--	--	--	--	AUC (AA) = .58 ²⁷ AUC (EA) = .64 AUC (NAH) = .73	
GRANN ET AL. (2000) ⁶	404	N/A	11.8 (3.7)	N/A	N/A	AUCs = .66 - .71	N/A
GRAY ET AL (2004)	315	19.9 (7.0)	11.4 (4.0)	3.8 (2.4)	4.7 (2.6)	AUCs = .61, .62, .48, .62 (Total, H, C, R)	N/A
GREVATT ET AL (2004)	44	19.4 (3.5) ²⁹	13.2 (3.2)	6.1 (2.0)	N/A	Inpat. AUCs = .56, .54, .60 (HC, C, R)	N/A
HILTERMAN ET AL (2002)	62	25.2 (7.5)	--	--	--	r_{tot} = -.37 w/BSI Direct Aggression Scale	IRR (TOTAL) = .98 (H,C,R,) = .92, .91, .95
MÜLLER-ISBERNER & JOCKEL (1997)	100	NA	11.5 (3.6)	5.2 (1.9)	NA	NA	M KAPPA (H) = .89 M KAPPA (C) = .49
MÜLLER-ISBERNER ET AL. (1999)	220	24.9 (5.9)	12.0 (3.4)	5.3 (2.2)	7.6 (1.9)	PEARSON R_{TOT} .20 - .40	KAPPA (HCR) = .72

CONTINUES ON NEXT PAGE WITH MORE FORENSIC AND CORRECTIONAL SAMPLES

STUDY/ SAMPLE	N	MEANS (SD)				VALIDITY INDICES	RELIABILITY INDICES
		HCR TOTAL	H SCALE	C SCALE	R SCALE		
FORENSIC PSYCHIATRIC							
NICHOLLS ET AL. (1999) ¹⁰	125	20.0 (5.3)	11.2 (3.6)	5.1 (2.5)	3.2 (1.2)	AUCs (TOT) = .68 - .77 PEARSON r = .31 - .46	N/A
PHAM (2001)	80	--	--	--		AUC (TOTAL) = .78	N/A
PHILIPSE (2002)	69	--	--	--	--	AUC (TOTAL) = .67 AUC (MODIFIED) ²¹ = .90	IRR (TOTAL) = .90
ROSS ET AL. (2001) ¹⁷	103	20.2 (5.6)	12.7 (3.5)	3.5 (2.1)	4.1 (2.3)	AUC (TOTAL) = .57; .76	N/A
SCHARIN (1999) ⁶	49	--	--	--	--	ODDS = 9.63	N/A
STRAND & BELFRAGE (2001)	63 ^B	24.8 (7.0) ^B	12.9 (3.6) ^B	5.1 (2.6) ^B	6.7 (2.9) ^B	NO DIFFERENCE B/W MEN AND WOMEN ON SCALES	KENDALL'S TAU-B = .67
	85 ^A	25.5 (7.9) ^A	13.8 (4.2) ^A	5.0 (2.5) ^A	6.7 (2.8) ^A		
STRAND ET AL. (1999)	40	26.4 (8.0)	14.4 (4.4)	5.2 (2.5)	6.8 (2.7)	AUC (TOTAL) = .80; COHEN'S <i>d</i> = 1.19	KENDALL'S TAU-B = .69
TENGSTRÖM (2001)	106	--		--	--	AUC (H) = .76	N/A
URHEIM ET AL (2003)	51	23.5 (6.8)	13.8 (4.3)	5.9 (1.9)	3.9 (2.0)	AUCs = .82, .77, .73, .76 ²⁹	
VINCENT (1998) ⁸	125	22.3 (6.3)	11.2 (3.6)	5.0 (2.6)	6.1 (2.1)	ODDS 2.45	N/A
WHITTEMORE (1999)	172					WALD = 9.86	
WINTRUP (1996) ⁴	80	NA ⁵	NA	NA	NA	M r = .30	NA
CORRECTIONAL							
BELFRAGE, FRANSSON, & STRAND (2000)	41	26				COHEN'S <i>d</i> = 1.70 FOR TOTAL, 1.00 FOR H, 1.14 FOR C, AND 1.22 FOR R	N/A

CONTINUES ON NEXT PAGE WITH MORE CORRECTIONAL SAMPLES

STUDY/ SAMPLE	N	MEANS (SD)				VALIDITY INDICES	RELIABILITY INDICES
		HCR TOTAL	H SCALE	C SCALE	R SCALE		
CORRECTIONAL							
COOKE ET AL. (2001) ¹⁴	250	?	10.9	?	?	AUC _{TOT} = .69 - .74 AUC _{VRAG} = .67 - .73 AUC _{PCL-R} = .65 - .72	ICC ₁ = .92, .92, .74, .70 (TOT, H, C, R)
DAHLE (2001)	200	19.0 (6.5)				r _{TOT} = .25	TAU _{TOT} = .80; RHO _H = .93; RHO _C = .73
DOUGLAS & WEBSTER (1999) ⁴	72	NA	11.9 (3.3)	5.0 (2.0)	NA	r _S = .3 - .5; M ODDS (HC) = 4.0	IRR (HC) = .80
DOUGLAS ET AL. (2003)	188	20.1 (7.9)	11.1 (3.9)	4.4 (2.5)	4.6 (2.8)	AUC (TOTAL) = .82 AUC (SPJ) = .78	IRR (TOTAL) = .93 IRR (SPJ) = .41
DOYLE ET AL. (2002)	87	NA	14.1 (3.9) ²² 11.4 (3.9)	NA	NA	AUCs (H) = .66 to .70	N/A
DUNBAR (2003)	58	22.2 (9.9)	9.4 (5.7)	6.1 (1.9)	6.7 (2.9)	rs (TOT) = .33 TO .63	IRR = .88 TO .94 α = .86 TO 94)
GRAY ET AL (2003)	34	NA	10.2 (3.4)	5.4 (2.6)	NA	INSTIT. AUCs = .81, .77, .79 (HC, H, C) INSTIT. rs = .53, .43, .49 (HC, H, C)	NA
KRONER & MILLS (2001)	97	17.8 (8.3)				INSTIT. rs = .11 TO .32 COMM. rs = .16 TO .39	ICC = .85 (Total)
PHAM ET AL (2000)	68	20.7 (9.2)	10.4 (5.0)	4.8 (2.4)	5.2 (2.4)	AUCs = .76, .77, .74, .71 (TOT, H, C, R)	r _{TOT} = .85
VINCENT (1998)	125	23.6 (6.7)	11.9 (3.8)	4.5 (2.5)	7.3 (1.7)	N/A	N/A
VINCENT ET AL. (2001) ¹⁵	56			3.9 (2.6) ^R 4.7 (2.6) ^C	5.8 (2.1) ^R 5.4 (2.6) ^C	N/A	ICC ₁ = .70 (C), .58 (R) ICC ₂ = .82 (C), .74 (R)

CONTINUES ON NEXT PAGE WITH MIXED AND JUVENILE SAMPLES

STUDY/ SAMPLE	N	MEANS (SD)				VALIDITY INDICES	RELIABILITY INDICES
		HCR TOTAL	H SCALE	C SCALE	R SCALE		
MIXED SAMPLES							
CÔTÉ (2001) ¹⁶	22 ^A	--	11.7 (4.3)	3.9 (2.2)	--	AUCs = .83, .61 (H, C)	ICC ₁ = .88 (H), .71 (C)
	36 ^B	--	13.1 (3.5)	4.8 (2.5)	--	Cohen's d = .29 – 1.4	ICC ₂ = .93 (H), .83 (C)
	19 ^C	--	16.0 (3.2)	4.3 (2.3)	--		
DOUGLAS & BELFRAGE (2002) ⁷						ds = (1) .89 - 1.75; (2) .36 - .50; (3) .08 - .44	
DOYLE & DOLAN (2003)	129	--	--	--	--	AUC (Total) = .62 - .80	N/A
FREESE ET AL. (2002)	128	22.2 (6.5) ²³	12.8 (3.7)	3.9 (2.8)	5.6 (2.2)	Cohen's d (Total) = .77	
		17.2 (6.5)	10.4 (4.4)	3.1 (1.9)	3.8 (2.2)		
HODGINS ET AL. (2001) ¹³	126					ICC ₁ = .85-.99 b/w raters and criteria scores	ICC ₁ = .90, .94, .89, .68 (Tot, H, C, R)
JUVENILE SAMPLES							
MACEachern (2001)	108	19.7 (6.6)	7.9 (3.2)	5.6 (1.9)	6.2 (2.3)	rs = .35 - .46 AUC = .73 - .79	ICC ₁ = .86, .88, .80, .77 (Tot, H, C, R)

Note 1. This table does not contain all studies reported in the Annotated bibliography. Some studies supplemental to main studies were not included. Other studies were excluded from the Table if they addressed issues other than the relationship between the HCR-20 and violence. The method and results of the studies in this table are described in more detail in the annotated bibliography that follows.

Note 2. IRR = Interrater Reliability; HC = Total H Scale and C Scale composite when R Scale not available; Superscript “A” denotes analyses for men only; Superscript “B” denotes analyses for women only. “ψ” denotes that the sample is a sub-sample of another study, and hence the reporting of *Ms* and *SDs* is omitted.

¹ See also Douglas, Ogloff, & Nicholls (1997a, b)

² See also Nicholls, Ogloff, & Douglas (1997a, b). Table reports validity indices for community violence only.

³ See also Klassen (1996)

⁴ See also Douglas, Webster, & Wintrup (1996)

⁵ The mean for the HC composite was 17.1 (*SD* = 3.8)

⁶ These samples are a combination of forensic and correctional.

⁷ These analyses are based on re-analysis of existing data sets across three samples, and hence *N*, descriptive characteristics, and reliability co-efficients are not provided here. the three d score ranges in the validity indices column refer to changes in C and R scores over time, and not to any relationship with violence.

⁸ The effect size was for predicting time institutionalized, not violence.

⁹ The effect size here was for predicting discharge from forensic hospital, not violence.

¹⁰ This sample also comprises the analyses for Vincent’s (1999) forensic sample.

¹¹ Eight patients were rated by 40 clinicians, where each patient was rated by five different clinicians

¹² The top mean (22.7) was derived from psychologists; the bottom mean (26.3) from psychiatric nurses.

¹³ There are 4 subsamples across different countries, with means broken down accordingly. They are not reported because of space.

¹⁴ Reported here for community violence only.

¹⁵ Means with superscript “R” refer to researcher-based ratings; those with superscript “C” refer to clinician-based ratings.

¹⁶ In Côté (2001), superscript “A” refers to involuntarily committed civil psychiatric patients, “B” refers to forensic patients (found not criminally responsible on account of mental disorder), and “C” refers to mentally disordered offenders.

¹⁷ Ross et al. (2001) overlaps in sample with Douglas et al. (1998).

¹⁸ Sample overlaps with Douglas et al. (1998); Ross et al (2001); Douglas & Ogloff (2003).

¹⁹ Means were provided for four different subgroups; range of means is presented.

²⁰ SPJ = Structured professional judgment of low, moderate, or high risk.

²¹ Based on a subset of most predictive items.

²² Larger mean is for the violent subgroup; smaller mean is for the nonviolent subgroup.

²³ Larger mean is for the violent subgroup; smaller mean is for the nonviolent subgroup.

²⁴ Psychopathy Item (H7) is omitted from mean values reported for HCR-20 Total and H-scale.

²⁵ AUCs were provided for readmission, self/collateral reports of violence, self/collateral reports of re-offending, and serious re-offending; range of AUCs is reported.

²⁶ Validity indices reported are for inpatient violence.

²⁷ Sample overlaps with Fujii, Lichten et al. (2004). Validity indices are for inpatient violence as a function of ethnic/racial group. AA = Asian American, EA = Euro-American, NAH = Native American part-Hawaiian

²⁸ Total Score is for HC composite.

²⁹ AUC values reported are for most severe episode of inpatient violence. The authors also presented AUCs for frequency.

OVERVIEW OF RESEARCH PROJECTS, PUBLICATIONS, PRESENTATIONS, AND UNPUBLISHED STUDIES

CIVIL PSYCHIATRIC SETTINGS

Project and Scholarly Work

McNiel, D., Gregory, A., Lam, J., Binder, R., & Sullivan, G. (2003). Utility of decision support tools for assessing acute risk of violence. *Journal of Consulting and Clinical Psychology, 71*, 945-953.

Summary

These authors used a pseudo-prospective design to evaluate the utility of three decision support tools for assessing acute risk of violence: the HCR-20, the PCL-SV, and the McNiel-Binder Violence Screening Checklist (VSC). 100 patients from a university-based, short-term psychiatric inpatient unit were used as participants. The design used a case-control method of sampling in which 50 individuals who had been physically assaultive were matched with 50 cases who had been nonviolent patients. For this study, the definition of violence was operationalized as physical attacks on persons. The median length of hospitalization was 9.5 days.

Inter-rater reliability as measured by ICC for the devices were: HCR-20 = .78, PCL-SV = .77, VSC = 1.0. The means from the study group were: HCR-20 total 18 ($SD = 6.6$), HCR-20 H-scale 7.1 ($SD = 3.5$), HCR-20 C-scale 6.1 ($SD = 2.3$) and HCR-20 R-scale 4.8 ($SD = 2.3$), VSC 2.1 ($SD = 1.3$), PCL-SV total score 9.1 ($SD = 5.1$), PCL-SV Part 1 4.7 ($SD = 3.0$) and Part 2 4.5 ($SD = 2.8$).

Correlational analyses showed that the HCR-20 total score was correlated with the PCL-SV total score ($r = .61$; $p < .01$) and with the VSC ($r = .26$; $p < .01$). Each of the HCR-20 scales is also correlated with the PCL-SV total score (H-scale ($r = .56$; $p < .01$), C-scale ($r = .4$; $p < .01$) and R-scale ($r = .47$; $p < .01$)) and with the VSC (H-scale ($r = .17$; $p < .01$), C-scale ($r = .34$; $p < .01$) and R-scale ($r = .15$; $p < .05$)).

Logistic regression analyses showed that when violence was predicted based on the total scores from the PCL-SV, the HCR-20 and the VSC, that only the VSC made an independent contribution to the violence prediction. Further regression analyses showed that when violence was predicted based on the subscale scores from the PCL-SV, the

subscales of the HCR-20 and the VSC, that the Clinical items from the HCR-20 and the VSC made independent contributions to violence prediction.

ROC analyses of the HCR-20 subscales showed AUC's of .56 for the H Scale, .77 for the C Scale and .58 for the R Scale. For the PCL:SV, the AUC for Part 1 was .66 and for Part 2 was .55. Of these subscales, only the HCR-20 Clinical items and Part 1 of the PCL-SV differed significantly ($p < .01$) from the line of no information. Compared to research using the HCR-20 with long-term community follow-up, the HCR-20 had generally lower levels of sensitivity and specificity in this sample.

The discussion section details the need for risk assessment tools as well the need for tools that are more appropriate for short-term risk assessment as opposed to long-term risk assessment. However, the C scale of the HCR-20 was shown to be an important independent predictor of short-term inpatient physical violence.

Project Description: Ogloff, J. R. P., Grant, I. An Investigation of Civil Commitment and Review Panel Decision Making in British Columbia

This was a chart review study of all 279 involuntarily committed persons from a large psychiatric hospital in Western Canada who applied for Review Panel hearings in 1994. Data were collected concerning patients' demographic characteristics, family and childhood history, mental health history, criminal history, and Review Panel hearing outcomes. The majority of patients had psychotic disorders, previous psychiatric hospitalizations, and were unemployed at admission. Over half of patients had previous arrests or convictions. Patients were tracked in the community after their release for an average of 2 years. Follow-up information was gathered from re-hospitalizations to the releasing psychiatric hospital, hospitalization records from 16 general hospitals in the province, provincial correctional records, and Coroner's records.

Scholarly Works

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

Summary

This study compared the predictive validity of the HCR-20 Risk Assessment Scheme (Webster, Douglas, Eaves, & Hart, 1997a; Webster, Eaves, Douglas, & Wintrup, 1995) and the Psychopathy Checklist: Screening Version (PCL:SV; Hart, Cox, & Hare, 1995). This research includes the 193 patients for whom complete measures were attainable (HCR-20; PCL:SV). Patients were followed into the community for an average of 626 days.

Violence was defined to include a demarcation between physical and non-physical aggression. Physical aggression refers to any attacks on persons. Non-physical aggression includes threats to harm a person, verbal attacks on persons, and "fear-inducing" behaviour such as attacks on objects. Violent crime was coded separately to allow for additional analyses, although typically it would also be coded as physical violence. The three types of violent outcome, then, were (1) any violence; (2) physical violence; (3) violent crime.

The AUCs produced by ROC ranged from .76 (for any and physical violence) to .80 (for violent crime). Odds ratios showed that persons scoring high on the HCR-20 (above the median) were 6 (for any and physical violence) to 13 (for violent crime) times more likely to be violent in the community than persons who scored under the median.

For the PCL:SV, AUCs ranged from .68 (for any violence) to .73 (for physical violence) to .79 (for violent crime). Effects for the PCL:SV were more variable than those for the HCR-20. The odds of violence for those above the median score of the PCL:SV also increased substantially (from approximately 4 to 13 times).

Hierarchical regression analyses revealed that the HCR-20 added to the predictive validity of the PCL:SV, but the converse was not true. Multiple regression analyses of the subscales of the HCR-20 and PCL:SV indicated that only HCR-20 scales predicted rate of violence. The H scale and R scale of the HCR-20 produced the largest effect sizes of all subscales with violence. Implications for research on risk assessment, as well as the clinical assessment and management of violence, are discussed.

Douglas, K. S., Ogloff, J. R. P., & Nicholls, T. L. (1997, June). *Assessing the risk for inpatient psychiatric violence*. Paper presented at the annual convention of the Canadian Psychological Association, Toronto.

Summary

This presentation focused on risk for inpatient violence specifically. Inpatient violence was defined in a similar manner as community violence. A distinction was made between physical violence (which required physical contact with victims) and non-physical violence (which included threats of violence and fear-inducing behaviour). Approximately half of patients displayed physical aggression while hospitalized.

AUCs for the H and C Scales composite for inpatient violence ranged from .57 to .65. Odds ratios for inpatient violence averaged approximately 2.0, and for repetitive inpatient violence, 3.0. These values are smaller than those for community violence, and indicate a moderately sized relationship between the HCR-20 and repetitive inpatient violence. The AUCs for the PCL:SV for inpatient violence were similar to those for the HCR-20, ranging from .60 to .64. Odds ratios were comparable to those of the HCR-20, averaging approximately 1.75 for inpatient violence, and 3.0 for repetitive violence.

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (2004). Assessing risk for violence among male and female psychiatric patients: The HCR-20, PCL:SV, and VSC. *Behavioral Sciences and the Law, 22*, 127-158.

Summary

The focus of this research was to compare the performance of the HCR-20, PCL:SV and a violence screening measure for both civil psychiatric inpatient and community violence. Concerning violent and criminal behaviour, a greater proportion of men had histories of crime (including violent crime). On the violence outcome measures, there were no differences in the incidence of inpatient violence across genders. A greater percentage of men compared to women displayed community violence.

Males had higher mean scores on the H Scale ($M = 10.8$; $SD = 3.3$), C Scale coded upon admission (7.4 ; $SD = 1.5$), and HCR-20 Total Scores ($M = 20.4$; $SD = 5.6$) compared to women (H Scale $M = 8.2$; $SD = 3.2$; C Scale at admission $M = 6.9$; $SD = 1.7$); HCR-20 Total Score $M = 16.8$; $SD = 5.4$). Males also had higher scores on the PCL:SV.

ROC AUC values for inpatient violence showed that the HC composite, the PCL:SV, and McNeil and Binder's (1994) Screening Measure did not predict violence for males. However, for females, moderate to large effects were observed for the HC composite ($AUCs = .62-.74$) and the PCL:SV ($AUCs = .63 = .74$). McNeil and Binder's screening measure was weakly related to violence, predicting only verbal aggression.

For post-release community violence, a different picture emerged. For male patients, HCR-20 AUCs ranged from .72 (any violence) to .73 (physical violence) to .75 (violence resulting in criminal sanctions). PCL:SV AUCs ranged from .63 (any violence) to .70 (violence resulting in criminal sanctions) to .71 (physical violence). For females, HCR-20 AUCs ranged from .66 (physical violence) to .77 (any violence) to .80 (violence resulting in criminal sanctions). PCL:SV AUCs ranged from .51 (physical violence) to .67 (any violence) to .89 (violence resulting in criminal sanctions).

This study is important because it focuses on gender. Perhaps surprisingly, it found that the HCR-20 performed better for the prediction of inpatient violence by women than by men. Prediction of community violence was comparable between genders. Statistical comparisons were not made between genders or measures, and as such the differential predictive validity was not addressed directly. Further, analyses were not carried out for HCR-20 and PCL:SV subscales.

See Also

Douglas, K. S., Ogloff, J. R. P., & Nicholls, T. L. (1997, August). *Violence by psychiatric patients: Validity of the HCR-20 Scheme and the Psychopathy Checklist: Screening Version*. Paper presented at the annual convention of the American Psychological Association, Chicago.

Douglas, K. S., Ogloff, J. R. P., & Nicholls, T. L. (1997, June). The role of personality disorders in community violence among civil psychiatric patients. In C. D. Webster (Symposium Moderator), *Personality disorder and violence*. Symposium presented at the Fifth International Congress of the Disorders of Personality, Vancouver, B.C., Canada.

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (1997, August). *Comparing risk assessments with female and male psychiatric outpatients: Utility of the HCR-20 and Psychopathy Checklist: Screening Version*. Paper presented at the annual convention of the American Psychological Association, Chicago.

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (1997, June). *Comparing risk assessments with female and male psychiatric inpatients: Utility of the HCR-20 and Psychopathy Checklist: Screening Version*. Paper presented at the annual convention of the Canadian Psychological Association, Toronto.

Ogloff, J. R. P., Douglas, K. S., Nicholls, T. N., & Grant, I. (1997, November). *Civil commitment and risk for violence in psychiatric patients*. Paper presented at the annual meeting of the Pinel Institute, Montreal, Quebec, Canada.

Ogloff, J. R. P., Nicholls, T. L., Douglas, K. S., & Grant, I. (1997, May). *Involuntary civil commitment: Risk assessment, sex differences, and review panel decision making*. Paper presented at the Annual Convention of the Law and Society Association, St. Louis, Missouri.

Project Description: Webster, C.D., Hart, S. D., Eaves, D. Prospective study of the HCR-20 in a civil psychiatric setting.

This was a prospective study of 131 persons admitted consecutively to the Intensive Care Unit (ICU) of a large psychiatric hospital in Western Canada. There were 82 (63%) men and 49 (37%) women. The mean age at admission was 36 years ($SD = 12$). The majority of patients were single ($n = 105$; 80%). Only 10% ($n = 13$) of the sample was employed at admission. The mean length of stay on the ICU was 21 days ($SD = 12$). Patients had on average 6.1 ($SD = 6.4$) previous psychiatric hospitalizations. Over half of the sample had schizophrenic or other psychotic disorders as admission diagnoses ($n = 73$; 56%). Approximately one-fifth ($n = 28$) of the sample received diagnoses of personality disorder.

The HCR-20, PCL:SV, and BPRS were completed for each patient. Research assistants coded the H scale items, and attending psychiatrists coded the C and R scale factors. Violence was measured on the unit by use of the Overt Aggression Scale. Patients were also tracked in the community. Subsequent contacts with corrections, police, and hospitals were recorded from archival sources. A research assistant also contacted community "collaterals" (persons who knew the patients and could report on their community behaviour) at three and six months post-release.

Scholarly Works

Klassen, C. (1996). *Predicting aggression in psychiatric inpatients using 10 historical risk factors: Validating the "H" of the HCR-20*. Unpublished bachelor's honour's thesis, Simon Fraser University, Burnaby, British Columbia, Canada.

Summary

In a subset of this sample comprising 50 patients, the 10 Historical variables of the HCR-20 and the 12 items from the PCL:SV were used to predict inpatient violence. Violence included acts of verbal aggression, self-directed aggression, and aggression toward others and objects (as measured by the Overt Aggression Scale). With respect to internal consistency of the HCR-20 H scale, Klassen reported a Cronbach's alpha of .73. Correlations between the H variables and violence averaged .30 across several out-

come measures, and controlling statistically for the effects of sex. Of the individual items, substance abuse and psychopathy were most strongly related to violence. The PCL:SV performed similarly to the H Scale, correlating at .26 with ward violence. Part 2 of the PCL:SV, which measures the behavioural aspects of psychopathy, was somewhat more strongly related to ward violence (.33) than were PCL:SV Total or H scores from the HCR-20.

Ross, D. J., Hart, S. D., & Webster, C. D. (1998). *Aggression in psychiatric patients: Using the HCR-20 to assess risk for violence in hospital and in the community*. Unpublished manuscript.

Summary

Interrater reliability for the H Scale, based on a subsample of 30 files, was .82. Cronbach's alpha for the H Scale was .74, and for the C Scale, .64. Interrater reliability for the PCL:SV Part 1, 2, and Total was, respectively, .82, .91, .91. In this study, 47% ($n = 62$) of patients displayed violence toward others while hospitalized. For inpatient violence, the H Scale, C Scale, and HC composite produced AUCs with violence that were greater than chance, ranging from .63 to .68 for any type of aggression. The largest AUC was for the HC composite. The PCL:SV AUC was .61. The HCR-20 H and C scales were related to ward violence with moderate strength in this sample. Survival analyses showed that persons who scored high on the HC composite were twice as likely (62%) to be violent by day 10 post-admission compared to persons who scored low (35%).

For the community phase of the study, 112 patients had been released by the end of the study period, and data were complete for 101 of these patients. Half of the sample displayed violent behaviour in the community, most frequently verbal aggression to others. For the HCR-20 subscales, AUCs for any aggression to others ranged from .58 (C), to .73 (R). For physical violence, the AUCs averaged approximately .63. The AUC for the HCR-20 Total score was .67. For violent crime, however, the HCR-20 AUC was .75. For the PCL:SV, the AUC for any violence and physical violence was .65, and for violent crime, .70. All AUCs are significantly greater than chance.

END OF CIVIL PSYCHIATRIC STUDIES

FORENSIC PSYCHIATRIC SETTINGS

Project and Scholarly Work

Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: Integrating research and clinical practice. *Journal of Forensic Psychiatry*, 9, 328-338.

Summary

This was a reliability study of the Swedish translation of the HCR-20 (Belfrage & Fransson, 1997). Six clinicians rated the same 43 patients on the HCR-20 and PCL-R. Over half of the sample ($n = 25$; 58%) had an index offence of homicide, and the majority (77%) had previous records for criminal offences. The mean age of patients was 40 (range = 24 - 67). The majority of patients had primary ICD-9 diagnoses of mental disorder (70%), 21% were diagnosed as personality disordered, and 9% received other diagnoses. Approximately half of the sample ($n = 22$; 51%) also had substance abuse diagnoses.

Internal consistency, using Cronbach's alpha, for the H scale was .96, for the C scale was .89, for the R scale was .85. For the total score, Cronbach's alpha was .95. Multivariate interrater reliability analyses, using Kendall's W , produced the following coefficients — Total scale = .81; H scale = .85; C scale = .62; R scale = .56. The HCR-20 correlated with the PCL-R at .64. The Cronbach's alpha for the PCL-R was .95, and Kendall's W was .78.

Project and Scholarly Work

Belfrage, H. & Douglas, K. S. (2002). Treatment effects on forensic psychiatric patients measured with the HCR-20 violence risk assessment scheme. *International Journal of Forensic Mental Health*, 1, 25-36.

Summary

This study used both cross-sectional and prospective methods to focus on the issue of change in HCR-20 violence risk factors in forensic psychiatric patients across multiple assessment periods. The sample consisted of 150 forensic psychiatric patients from two maximum security forensic psychiatric hospitals in Sweden. The sample was all male, the majority had committed violent crimes (94%) and had been assessed on more than one occasion. For the cross-sectional analyses, the sample was divided into three

groups: those who had been institutionalized up to a year, between one and two years, and more than two years. A sub-sample of 70 men was followed prospectively across three assessment periods with six months in between each assessment to further analyze change in violence risk factors.

Cross-sectional results showed that the mean scores for the C-scale and the R-scale of the HCR-20 were significantly lower the longer that patients had been hospitalized. These results were only significant for the C-scale when comparing the group which had been institutionalized for up to one year against those who had been in for over two years (C-scale $p \leq .038$). The R-scale showed significant changes between the one year group and the one-two year group ($p = .01$) as well as between the one year group and the more than two years group ($p < .001$).

The within-groups prospective analyses contained 70 subjects whose treatment times were much longer than those in the previous analyses. For this group, the mean scores from the C-scale dropped significantly over time both between time 1 and time 2 ($t = 2.07$; $p < .05$) and between time 1 and time 3 ($t = 2.96$; $p < .01$). However, the scores from the R-scale did not drop significantly for either time period.

Project and Scholarly Work

Brown, L. K. (2001). *Assessing risk for elopement and breaches of conditional release in insanity acquittees*. Unpublished Ph.D. Dissertation. Simon Fraser University, Burnaby, British Columbia, Canada.

Summary

This study evaluated the utility of the HCR-20, PCL-R, and VRAG in predicting negative outcomes of people found NCRMD. The sample comprised 172 insanity acquittees (20 women and 152 men) appearing before a criminal Review Board in British Columbia. Participants' mean age was 34.17 years ($SD = 9.70$). Most (91%) had primary diagnoses of a psychotic disorder (6% organic mental disorder; 2% anxiety or other disorder; 1% substance abuse disorder). Almost half (42%) had secondary diagnoses of substance abuse or dependence and 31% were diagnosed with a personality disorder.

The C and R scales were completed by psychiatrists as part of their routine assessment prior to the Review Board Hearing. The H scale was completed by a research assistant using file material. For 67 participants, the PCL-R was completed using both an interview and file material; for the remainder of the sample, only file material was used. All instruments were completed prospectively except for

the VRAG, which was coded at the end of the study period.

The mean PCL-R scores were: total = 16.51 ($SD = 7.27$); F1 = 5.92 ($SD = 3.14$); F2 = 8.84 ($SD = 4.00$). The mean VRAG score was 0.58 ($SD = 8.92$). The mean HCR-20 scores were: total = 22.67 ($SD = 6.53$); H scale = 12.80 ($SD = 3.64$); C scale = 4.47 ($SD = 2.54$); R scale = 5.41 ($SD = 2.71$).

Elopers had significantly higher scores on the H scale compared to non-elothers; $t(109.62) = 3.58, p = .001$. The two groups did not have significantly different scores on the C ($p = .548$) and R ($p = .342$) scales. Elothers also had significantly higher PCL-R scores and were placed into higher VRAG score bins relative to non-elothers. Cox regression analyses using the HCR-20 scales, with time at risk calculated as time spent in hospital during the study period, yielded the same pattern of results as the univariate analyses. When the individual HCR-20 items were examined, only Item H10 (prior supervision failure) was positively and significantly associated with risk of elopement. When HCR-20 scales were compared to PCL-R total scores and VRAG bin scores controlling for age, none of the variables was related significantly to risk of elopement.

There were 109 participants who were released on conditional discharge. Compared to participants not released, those who were released had significantly lower mean scores on the C scale ($t = 6.74, p \leq .001$) and R scale ($t = 9.61, p \leq .001$). The groups did not have significantly different H scale ($p = .843$), PCL-R ($p = .603$), or VRAG bin scores ($p = .790$). Cox regression analyses indicated that R scale scores were associated with likelihood of release (Wald = 23.06, $p \leq .001$), but H (Wald = .42, $p = .517$) and C scale (Wald = .36, $p = .550$) scores were not. When individual HCR-20 items were considered, negative and significant associations with release were found for previous violence, active symptoms of major mental illness, and plans lack feasibility. In another Cox regression analysis that compared the three HCR-20 scales, PCL-R, and VRAG bin scores controlling for age, a significant (positive) relation was found only for the R scale. Age was associated negatively with release.

Of the 109 participants released, 43 were returned and one committed suicide. The following analyses consider only the first rehospitalization in those cases where there were multiple returns for the same individual. Neither univariate analyses nor Cox regression analyses indicated significant differences on any of the measures between those who were successful or failed on release. When the individual HCR-20 items were considered, a positive and significant relationship was found only for active symptoms of major mental illness. When the dependant variable was narrowed to rehospitalization following a significant security problem in the community, PCL-R (Wald = 9.41; $p = .002$) and

R (Wald = 3.89; $p = .049$) scale scores were significant positive predictors. The H scale was significantly but negatively related to this return following a security problem (Wald = 6.89, $p = .009$).

Project and Scholarly Work

Claix, A., Pham, T., & Willocq, L. (2002, March). *Evaluation of the HCR-20 (Historical-Clinical-Risk management) in a Belgian forensic population*. Poster presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.

Summary

This study reported on the descriptive statistics of the HCR-20 as well as assessing the relations between the HCR-20, the PCL-R, and the Buss and Perry Aggression Questionnaire (Buss and Perry, 1992). This study used 86 French speaking male adult offenders confined in a Belgian forensic hospital.

Types of offenses measured were: violent offenses, non-violent offenses and any sex offenses. Common items between the HCR-20 and the PCL-R and the AQ were omitted. The omitted items were H7 (psychopathy), C1 (introspection) and H1 (past violent behavior) from the HCR-20.

The HCR-20 total score had adequate inter-rater reliability ($r = .73$) and good internal consistency (Cronbach's alpha = .74). The inter-rater reliability for the H-scale alone was ($r = .85; p < .01$) with an internal consistency alpha of .61. The inter-rater reliability for the C-scale alone was ($r = .65; p < .05$) with an internal consistency alpha of .47. The inter-rater reliability for the R-scale alone was ($r = .64; p < .05$) with an internal consistency alpha of .54.

The HCR-20 and the PCL-R were highly correlated across most of their scales. The total, H and C scales from the HCR-20 were all significantly ($p < .01$) and highly (r 's $> .4$) correlated with the PCL-R total, Factor 1 and Factor 2 scales. The HCR-20 R scale was only correlated at the $p < .05$ level and with r 's between .22 and .25 with the PCL-R scales. Using only a sub sample of 70 men, the HCR-20 scales showed far fewer significant correlations with the AQ. The HCR-20 total score ($r = .3; p < .05$) and the H-scale score ($r = .39; p < .01$) were correlated with the AQ total score. The HCR-20 total score ($r = .34; p < .01$) and the H-scale score ($r = .46; p < .01$) were also correlated with the AQ physical aggression score. The other HCR-20 scales were not significantly related to the AQ scores.

The HCR-20 scores were correlated to a few types of violent offenses. The HCR-20 total score was correlated with violent theft ($r = .26; p < .05$) and with assault and battery

($r = .3$; $p < .01$). The H scale was also correlated with violent theft ($r = .26$; $p < .05$) and with assault and battery ($r = .37$; $p < .01$). The C scale was correlated with kidnapping ($r = .26$; $p < .05$).

The HCR-20 scores were correlated to a few types of non-violent offenses. The HCR-20 total score was correlated with theft ($r = .28$; $p < .01$). The H scale was also correlated with theft ($r = .27$; $p < .05$) as well as drug offenses ($r = .24$; $p < .05$).

Lastly, the HCR-20 scores were correlated with indices of homicide. The HCR-20 total score was correlated with psychotic homicide ($r = -.74$; $p < .01$). The H scale was also correlated with psychotic homicide ($r = -.67$; $p < .05$). The C scale was also correlated with psychotic homicide ($r = -.64$; $p < .05$) and reactive homicide ($r = -.56$; $p < .05$) and with instrumental homicide ($r = .71$; $p < .01$).

Project and Scholarly Work

Dernevik, M. (1998). Preliminary findings on reliability and validity of the Historical-Clinical-Risk Assessment in a forensic psychiatric setting. *Psychology, Crime, and Law*, 4, 127-137.

Summary

This was a reliability study of the HCR-20. Six clinicians each rated six patients on the HCR-20. Reliability coefficients ranged from .76 to .96.

Project and Scholarly Work

Dernevik, M., Falkheim, M., Holmqvist, R. & Sandell, R. (2001) Implementing Risk Assessment: Clinical Judgment Revisited. In D. Farrington, C. Hollin, & M. McMurrin (Eds.). *Sex and Violence: The Psychology of Crime and Risk Assessment*, Harwood Academic Press. London.

Summary

The main goal of this study was to evaluate issues related to the process of risk assessment as it pertains to the HCR-20. Specifically, the study evaluated whether “expert” HCR-20 raters (psychologists) differed in their scores from psychiatric nurses. Second, analyses were conducted to determine the extent to which HCR-20 ratings were influenced by clinicians’ feelings towards the patient. The contextual grounding for this approach was drawn from the larger clinical and social psychological literature on biases and heuristics in decision-making.

A total of 8 male patients and 40 clinicians (psychiatric nurses) took part. On average, each patient was rated by five clinicians, and each clinician rated one patient. These patients had serious violent index offences (homicide, rape, assault, arson) and severe mental disorders, as well as personality disorders. They were on average 28 years of age.

The “Feeling Word Checklist” (FWC) was used for clinicians to rate their reactions to the patients they assessed. The FWC is based on a circumplex model with 30 items comprising four dimensions and eight scales, as follows: (1) Helpfulness vs. Unhelpfulness; (2) Closeness vs. Distance; (3) Accepting vs. Rejecting; and (4) Autonomous vs. Rejecting. The FWC predicted HCR-20 scores with $Mult. R = .66$, with feeling Close and Accepting relating to higher scores, and Helpfulness and Autonomy relating to lower scores.

The mean score for the nurse was 26.3 ($SD = 6.1$), whereas it was lower for the “expert” raters ($M = 22.7$; $SD = 6.5$).

As Dernevik et al. point out, the question of whether the relationship between feelings and HCR-20 scores is evidence for biases in clinical decision-making is not clear. There were no outcome data (i.e., subsequent violence). Further, it is possible that clinicians’ feelings are correlative rather than causative of the HCR-20 ratings. Dernevik et al.’s findings, however, emphasize the importance of limiting biases and over-emphasis on personality to the greatest extent possible, and also the potential importance in professional training on the outcome of an assessment. Further, item bias was not directly assessed (i.e., differential item functioning using Item Response Theory).

See Also

Dernevik, M., Falkheim, M., Holmqvist, R., & Sandell, R. (1999, July). Implementing risk assessment procedures in a forensic psychiatric setting: Personal relationships between assessor and the assessed using the Historical-Clinical-Risk-20 scheme. Paper presented at the International meeting of the American Psychology-Law Society (Div. 41 APA) and the European Academy of Psychology and Law, Dublin.

Project and Scholarly Work

Dernevik, M., Grann, M., & Johansson, S. (2002). Violent behaviour in forensic psychiatric patients: Risk assessment and different risk management levels using the HCR-20. *Psychology, Crime, and Law*, 8, 83-111.

Summary

This was a prospective study of short-term inpatient and community violence. Participants were 54 consecutive admissions over two years to a forensic psychiatric unit. Most ($n = 48$) were male. Mean age was 34.2 ($SD = 8.92$). Most had violent index offences (assault, $n = 16$; murder, $n = 10$; great bodily harm, $n = 4$; arson, $n = 10$; sex offences, $n = 6$; other, $n = 8$). 29% had an Axis I diagnosis only (mostly schizophrenia); 14% had Axis II only; 27% had both; 9% had other combinations of diagnoses.

Predictive analyses were carried out for the whole sample, as well as across three risk management levels: *Level one:* (High RM) Time spent on a high security ward with no access to the community. *Level two:* (Medium RM) Time spent living in the hospital but with limited access to occupational and recreational activities in the community. *Level three:* (Low RM) Time spent in less secure living arrangements and having access to the community while still being monitored regularly.

For overall analyses, HCR-20 effects with inpatient violence were as follows: HCR-20 Total Score ($r = .32$; $AUC = .68$); H Scale ($r = .37$; $AUC = .68$); C, R, and PCL:SV did not predict inpatient violence. For community violence re-conviction analyses, HCR-20 Total Score $AUC = .84$; PCL:SV $AUC = .71$. The C Scale had the highest AUC of the subscales, at .79.

In the low, medium, and high risk management conditions, the measures were most predictive in low and medium conditions, and less so in the high risk management condition. In the High Risk Management condition, only the H Scale was predictive ($AUC = .67$). HCR-20 Total Score predicted with $r = .21$ and $AUC = .64$. C, R, and PCL:SV did not predict. In the Medium Risk Management condition, effects were as follows: HCR-20 Total Score ($r = .41$; $AUC = .82$); H Scale ($r = .34$; $AUC = .83$); C Scale ($r = .36$; $AUC = .75$). R and PCL:SV were not significantly associated with violence, though had small/moderate effect sizes. In the Low Risk Management condition, HCR-20 Total Score ($r = .50$; $AUC = .71$); H Scale ($r = .48$; $AUC = .75$); R ($r = .49$; $AUC = .62$); C and PCL:SV did not predict.

Dernevik et al. interpreted their results as supporting the predictive validity of the HCR-20 for inpatient and community violence. The finding that the HCR-20 was less strongly related to violence in the High Risk Management than in the Medium or Low Risk Management categories, or in the community follow-up, was interpreted not as lack of validity but as effective intensive clinical risk management in this category. This is consistent with the finding and conclusion reached by Muller-Isberner et al. (1999).

Project and Scholarly Work

Dolan, M., & Khawaja, A. (2004). The HCR-20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior*, 30, 469-483.

Summary

This study investigated the predictive validity of the HCR-20 total and subscale scores among violent patients ($N = 70$) discharged to the community under fairly intensive supervision. Participants were discharged between 1992 and 2000 and had stayed in the hospital for a mean of 24 months ($SD = 14.49$). Most of the sample was Caucasian (83%) and single (80%). The mean age at admission was 35.3 years ($SD = 10.12$). The most common Axis I diagnosis was schizophrenia (73%) and roughly one-third of the sample had either primary or secondary diagnoses of a PD (the most common being APD, 26%). Almost half (44%) had a history of co-morbid substance misuse.

The HCR-20 was scored from comprehensive case file information at the time of discharge. Item H7 (PCL-R score) was not coded for this study. Three types of follow-up outcome data were collected blind to the initial HCR-20 scores: (1) reconvictions were coded from the Home Office Offender Index; (2) readmissions to district and forensic hospitals (readmissions could be of several types, including those under the Mental Health Act that reflected concern over an escalation in violence secondary to a relapse in mental state); and (3) self/collateral reports of violence were coded from community mental health teams' computerized records. Violence included sexual violence, punching, biting, choking, kicking, or assault with a weapon that resulted in physical injury to the victim.

Mean HCR-20 scores, with the psychopathy item omitted, were: total = 19.37 ($SD = 5.7$, range = 9-31); H-scale = 11.82 ($SD = 3.65$, range = 6-18); C-scale = 3.34 ($SD = 2.20$, range = 0-10); R-scale = 4.07 ($SD = 1.45$, range = 1-7).

The mean length of stay in the community was 59 months ($SD = 37.64$). There was a significant negative relationship between time in the community and HCR-20 total score ($r = -.48$, $p = .0001$), H-scale ($r = -.29$, $p = .014$), C-scale ($r = -.57$, $p = .001$), and R-scale ($r = -.37$, $p = .001$).

For the following analyses, median splits were performed and high and low scores refer to scores above and below the median, respectively. No significant associations were detected between high and low total scores on the HCR-20 and re-offending ($\chi^2 = 2.71$, $p = .10$) or violent re-offending ($\chi^2 = 1.72$, $p = .18$). Likewise, none of the subscales demonstrated a significant association with reconviction. However, the number of readmissions was correlated significantly with HCR-20 total score ($r = .40$, $p =$

.0001), C-scale ($r = .26, p = .026$), and R-scale ($r = .31, p = .007$), but not H-scale. Chi-square analyses indicated a significant association between high total scores and all subscales for any form of readmission and especially for readmission under the Mental Health Act (MHA). For incidents of violence reported by participants or their carers, there was a significant association with high HCR-20 total scores ($\chi^2 = 10.19, p = .006$), H-scale ($\chi^2 = 16.13, p = .001$), and C-scale ($\chi^2 = 7.46, p = .008$). This association nearly reached significance for the R-scale ($\chi^2 = 3.06, p = .08$).

ROC curve analyses were used as another index of predictive validity. The AUC for the HCR-20 total score for re-admission under the MHA was .85 ($p < .001, CI = .76 - .95$). AUC values also were significant for self/collateral reports of violence (AUC = .76, $p < .001, CI = .65-.87$) and re-offending (AUC = .71, $p < .05, CI = .56-.87$). The AUC value for serious re-offending was not significant (AUC = .67, $p = .15, CI = .47-.88$). All three scales predicted readmission under the MHA (AUC values ranging from .65 to .78), with highest values obtained for the H-scale.

Kaplan-Meier survival analyses revealed significant relationships between above-median HCR-20 scores and poor survival in the community. Log rank values were: MHA readmission = 27.73 ($p < .001$); self/collateral reported violence = 17.14 ($p < .001$); re-offending = 3.85 ($p < .05$); and violent re-offending = 5.08 ($p < .05$). The authors noted (but did not provide a quantitative summary) that when survival analyses were used to examine time at risk in the community, they found evidence that the C-scale and R-scale outperformed the H-scale.

The discussion section highlights reasons that may have contributed to the finding of a significant relationship between HCR-20 scores and readmission but not reconviction (e.g., high level of supervision, hospital policy pertaining to re-admittance at time of deterioration in mental state/increase in risk of violence).

Project Description: Webster, C. D., Hart, S. D., & Eaves, D. Prospective study of the HCR-20 in a forensic psychiatric setting.

This is a prospective study. The HCR-20 was coded on 175 consecutive persons who were coming before a Criminal Review Board for release from dispositions of Not Criminally Responsible an Account of Mental Disorder (NCRMD). The PCL-R was coded with the use of interviews by trained assistants. Psychiatrists who were responsible for providing the Board with a release assessment completed the Brief Psychiatric Rating Scale

(BPRS), as well as the Clinical and Risk Management scales of the HCR-20, as part of their assessments. The H scale was coded by assistants on the basis of file and interview information.

The mean age at admission was 33.0 ($SD = 9.6$). The sample was primarily male ($n = 133, 88.7\%$). The vast majority of participants were unemployed at admission ($n = 139, 92.7\%$). Schizophrenia was the predominant admission Axis I diagnosis ($n = 96, 64\%$). Forty-one patients (27.4%) of the patients received an admission diagnosis of personality disorder. Most patients had been hospitalized in the past ($n = 132, 88\%$), and the majority had previous charges for violent offences ($n = 90, 60\%$). Finally, most patients had a violent index offence ($n = 129, 86\%$). Violence was measured in the hospital with the Overt Aggression Scale, and in the community with arrests records and re-admission to the forensic institute.

Scholarly Works

Douglas, K. S., Klassen, C., Ross, D., Hart, S. D., Webster, C. D., & Eaves, D. (1998, August). *Psychometric properties of HCR-20 violence risk assessment scheme in insanity acquittees*. Poster presented at the Annual meeting of the American Psychological Association, 1998, San Francisco.

Summary

The HCR-20 violence risk assessment scheme was coded on 175 consecutive insanity acquittees appearing before a criminal Review Board. The purpose of the study was to provide data on the descriptive, normative, and reliability characteristics of the HCR-20, and on its relationship to conceptually-related concurrent measures and indexes. The alpha co-efficients for the HCR-20 Total, H scale, C scale, and R scale scores, respectively, were .78, .69, .77, and .77. Other indexes also supported the structural reliability of the HCR-20 (i.e., *MIC*; *CITC*). For the H Scale, inter-rater reliability was good ($ICC_1 = .81; ICC_2 = .90$). Inter-rater reliability was not available for the other HCR-20 scales. Test-retest analyses showed that the C and R scales changed (declined) across repeated assessments, as they are expected to.

The HCR-20 was related strongly to the PCL-R, correlating at .60. The H Scale was most strongly related (.76 with PCL-R Total), while the C and R Scales were related with small effect sizes ($r_s = .18$ and $.16$, respectively). The H Scale was more strongly correlated with Factor 2 of the PCL-R, while the C and R Scales were more strongly correlated with Factor 1. The HCR-20 and its scales were related to psychopathology (Brief Psychiatric Rating Scale; various factors). In a conceptually meaningful way. Gener-

ally, the C Scale was most strongly related, the R Scale next strongly related, and the H Scale generally unrelated.

Finally, the HCR-20 was related to an index of violence (past violent crimes). Items on the HCR-20 dealing with past violence were removed to avoid inflation of correlation co-efficients. Persons scoring above the median of the HCR-20 were significantly more likely than those scoring below the median to have previous violent convictions, previous assault charges, and juvenile records.

Douglas, K. S., & Ogloff, J. R. P. (2003). Multiple facets of risk for violence: The impact of judgmental specificity on structured decisions about violence risk. *International Journal of Forensic Mental Health, 1*, 19-34.

Summary

The conceptual risk assessment literature describes risk as multi-faceted (i.e., likelihood; severity; imminence; nature; targets) and calls for decision-makers to make decisions about these various facets. However, no research has been conducted to evaluate whether such highly specific judgments can be made reliably or accurately. This study evaluated whether highly specific judgments of violence could be made with reasonable reliability and incremental validity over more general judgments. For this study, the authors used the HCR-20 as the primary measure of violence risk. The sample consisted of 100 adults who had been found not criminally responsible by reason of insanity and were released from a maximum-security forensic institution into the community in 1996. The HCR-20 (Version) 2 was recoded from the original 175 participants described above because Version 1 had originally been used. The majority of the sample was male ($n = 89$). For this study, the definition of violence was categorized into three groups: any violence, physical violence, non-physical violence, and violence that resulted in criminal charges. Violence was measured from two sources: official criminal recidivism data and records of readmission to forensic psychiatric services.

The inter-rater reliability was calculated using ICCs. For the omnibus (general) structured clinical risk ratings on the HCR-20, the ICC_1 was .61. Reliability was lower for more specific judgments – from low to moderate for judgments of various severities of violence ($ICC_1 = .27-.37$). Reliability was low to moderate for ratings of violence targets ($ICC_1 = .40-.47$). Lastly, reliability was low to moderate for time frame of violence of up to one year ($ICC_1 = .31-.42$).

Only a limited number of the specific judgments about future violence were capable of being evaluated due to low base rates, low reliability or lack of ability to collect outcome information. The short-term risk judgment produced

a small but significant point biserial correlation with violence at 12 months ($r = .2; p = .02$). A partial point biserial correlation between short-term risk judgments and 12-month violence holding the general risk assessment constant was .06. Risk judgments of minor violence showed a significant point biserial correlation with non-physical violence ($r = .23; p = .012$). The partial point biserial correlation holding omnibus risk judgment constant was non-significant. Risk judgments of severe violence showed a significant point biserial correlation with physical violence ($r = .27; p = .003$). The partial point biserial correlation holding omnibus risk judgment constant was non-significant.

The findings showed that more specific judgments of various facets of violence risk that are called for in the literature were not made with as much reliability and accuracy as more general judgments.

Douglas, K. S., & Ogloff, J. R. P. (2003). The impact of confidence on the accuracy of structured professional and actuarial violence risk judgments in a sample of forensic psychiatric patients. *Law and Human Behavior, 27*, 573-587.

Summary

The purpose of the study was to evaluate the relationship between confidence and accuracy of risk estimates made using a structured professional judgment (SPJ) and an actuarial approach. The impetus for the study was previous research (McNiel, Sandberg, & Binder, 1998) in which probabilistic clinical predictions of inpatient violence by civil psychiatric patients were influenced by clinicians' confidence in their judgments.

The sample comprised 100 forensic psychiatric patients in western North America who had been found to be not criminally responsible for criminal offenses (previously reported on by Douglas, Ogloff, & Hart, 2003). The mean age at admission was 35.30 years ($SD = 9.84$) and participants primarily were single (67%), unemployed (93%), and had an admission diagnosis of schizophrenia (73.5%); 24.0% personality disorder; 18.4% mood disorder; 5.1% substance use/abuse disorder; 3% 'other'). Most participants had a past violent charge (91.9%) and almost half had a past violent conviction (48.5%). The majority (79.0%) had a violent index offense.

The HCR-20 was coded archivally by two masters-level clinicians who were blind to outcome. Raters made actuarial predictions of risk (the sum of the HCR-20 items for each scale) and SPJ predictions of risk (final risk judgments of low, moderate, or high risk). They rated their confidence in their HCR-20 judgments on a 1-10 scale, where confidence was defined as "the rater has a feeling of

certainty or reliance or trust about the correctness of the rating.” A median split was applied to create a low confidence group (who scored at or below the median) and a high confidence group (who scored above the median).

Four categories of violence were coded from criminal and hospital readmission records: (1) physical violence (physical contact by the perpetrator or use of a weapon); (2) non-physical violence (verbal threats and fear-inducing behavior); (3) criminal violence (violence that led to arrests or convictions); and (4) any violence (an omnibus category that included all violence).

A striking contrast emerged between effects of SPJs across the high and low confidence groups, with point biserial correlations (r_{pb}) and AUCs in the former typically being large and significant but in the latter being not significant. Correlations for any, physical, nonphysical, and criminal violence for the high confidence group were .62, .54, .48, and .43 and for the low confidence group were .14, .18, .10, and .03, respectively. AUC values for the any, physical, nonphysical, and criminal violence for the high confidence group were .86, .82, .82, and .84 and for the low confidence group were .58, .63, .58, and .52, respectively. Cox regression analyses, which control for time and uneven follow-up periods, yielded a nonsignificant model fit for the low confidence group using ‘any violence’ as the outcome criteria. However, in the high confidence group there was a roughly ninefold increase in the hazard of violence that occurred between low and moderate and between moderate and high risk ratings.

A similar set of analyses was carried out for the three actuarial judgments (one for each scale). In the low confidence group, all r_{pb} and AUC values across the four violence categories were nonsignificant and generally small, whereas for the high confidence group the values generally were larger and were significant for the H- and C-scales (but not for the R-scale).

Indices of variability for scale scores and final risk ratings were highly comparable between the high and low confidence groups, which provided evidence against the possibility that the results could be attributed to differential variance of the predictors between the two confidence groups. Several possible explanations for the strong relationship observed between confidence and accuracy are discussed.

Douglas, K. S., Ogloff, J. R. P., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services, 54*, 1372-1379.

Summary

This study tested the inter-rater reliability and criterion-related validity of structured violence risk judgments made with one application of the structured professional judgment (SPJ) model of violence risk assessment, the HCR-20 violence risk assessment scheme. Participants were taken from a larger, ongoing prospective study examining the predictive validity of the HCR-20. From the larger study, 116 of 175 patients released from forensic hospitalization between 1996-1997 were originally chosen to participate. The HCR-20 was completed on a random sample of 100 of the 116 forensic psychiatric patients. All of the 100 had been found not guilty by reason of insanity and were subsequently released into the community.

For this study, violence was operationally defined as actual, attempted or threatened physical harm to others. Acts of violence were divided into broad categories of: any violence, physical violence and non-physical violence. Raters were two masters-level clinicians. Raters gathered information from clinical-legal files of participants as they existed at time of discharge. Violence in the community was coded both from criminal records and clinical files after discharge from the hospital.

The mean HCR-20 total score was 24.7 ($SD = 4.64$). For the H-scale the mean was 14.4 ($SD = 2.79$), for the C-scale it was 4.68 ($SD=2.02$) and for the R-scale it was 5.88 ($SD = 1.49$). The *ICC* for the H-scale ranged from .41 (H4) to 1.0 (H7). For the total H scale it was .90. The *ICC* for the C-scale ranged from .34 (C5) to .69 (C3) (total C scale = .79), and for the R-scale, the *ICC* ranged from .01 (R5) to .54 (R3) (total R scale = .47). *ICC* for the HCR-20 total score was .85. As for agreement on final risk ratings, raters agreed on 70% of all cases, with no instances of low/high risk errors ($ICC = .61$).

AUC values for the HCR-20 structured clinical judgments (low, moderate, or high risk) were statistically significant for each outcome criterion. Effects for the HCR-20 clinical judgments were moderate to large in size, depending on the violence index (any violence, $AUC = .69, p < .01$; physical violence, $AUC = .74, p < .01$; non-physical violence, $AUC = .68, p < .05$). For the HCR-20 total score, the *AUC* for any violence was .67, $p < .05$; for physical violence was .70, $p < .05$ and for non-physical violence was .67, $p < .05$. For the H-scale, the *AUC*s were not significant. For the C-scale, the *AUC* for any violence was .68, $p < .05$; for physical violence it was .70, $p < .05$ and for non-physical violence it was .68, $p < .05$. For the R-scale, the *AUC*'s were not significant.

Kaplan-Meier survival analyses showed that persons judged to be at high risk were more likely to be violent, and to be so sooner than others. Cox regression analyses showed that HCR-20 risk ratings were most strongly related to violence, over and above actuarial scores.

The discussion section reiterates the findings and explores the implications of these results for using structured clinical judgments in risk assessments.

Ross, D. J., Hart, S. D., Eaves, D., & Webster, C. D. (1999, November). *The relationship between the HCR-20 and BC Review Board decisions on the release of forensic psychiatric inpatients*. Paper presented at the International Conference on Risk Assessment and Risk Management, Vancouver, British Columbia, Canada.

Summary

This presentation compared the H, C, and R scores of those patients who had been released by the Review Board to those who had not. While the H scale score did not differ between groups, C and R scale scores did. Among those discharged, the C scale score was 3.4, compared to 5.9 among those not released. Similarly, the R scale score was significantly lower among those released (4.0) compared to those not released (7.3).

Ross, D. J., Hart, S. D., Eaves, D., & Webster, C. D. (2001, April). *The relationship between the HCR-20 and community violence in a sample of NCRMD outpatients*. Paper presented at the Founding Conference of the International Association of Forensic Mental Health Services, Vancouver, British Columbia, Canada

Summary

This presentation reported the results of prospective analyses of the prediction of post-release violence among 103 released forensic patients followed for six months. The AUC value between “any aggression” and total score was .76. For H, C, and R, it was .60, .74, and .75. AUC values for PCL:SV total, Part 1, and Part 2 were .64, .57, and .66. For physical aggression, the AUC values were smaller: .57, .57, .60, and .61 for HCR-20 total, H, C, and R scale scores. They were larger for PCL:SV total, Part 1, and Part 2 scores: .77, .75, and .70. As with Dernevik et al. (2002) and Müller-Isberner et al. (1999), Ross et al. (2001) suggested that risk management strategies could be responsible for the lower effects observed for more serious violence and HCR-20 scores. As with the other studies, however, this hypothesis remains untested. It is important to point out that the lower effects for more serious violence does not necessarily reflect a trend across studies, as other reports have failed to observe this (Douglas et al., 1999).

Project and Scholarly Work

de Vogel, V. & de Ruiter, C. (in press). The HCR-20 in personality disordered female offenders: A comparison with a matched sample of males. *Clinical Psychology and Psychotherapy*.

Summary

The predictive validity of the Dutch version of the HCR-20 was examined in a forensic psychiatric sample of 42 women admitted between 1985 and 2003. A sample of 42 male forensic psychiatric patients, also admitted between 1985 and 2003 and matched on birth year, type of index offense, ethnicity, and type of psychopathology, was used as a comparison group.

The HCR-20 was coded on the basis of file information. For the women, ratings were made retrospectively for 15 cases and prospectively for 27 cases. Good interrater reliability was observed for the women for the total score, H-scale, and final risk judgment ($n = 27$; ICCs = .75, .82, .74) and moderate for the C-scale and R-scale (ICCs = .55, .51). For the men, half the ratings were retrospective and half were prospective. Good interrater reliability was observed for the total score, H-scale, C-scale, and final risk judgment ($n = 28$; ICCs = .77, .82, .70, .69).

There were significant mean differences between the genders on several HCR-20 items but the total and scale scores were comparable. For women, mean scores were: HCR-20 total (25.9, $SD = 5.5$); H-scale (14.0, $SD = 2.9$); C-scale (5.4, $SD = 2.0$); R-scale (6.6, $SD = 1.9$). For men, mean scores were: HCR-20 total (27.1, $SD = 6.5$); H-scale (14.9, $SD = 3.0$); C-scale (5.4, $SD = 2.3$); R-scale (6.8, $SD = 2.1$). With respect to the HCR-20 final risk judgments, women were judged as moderate risk significantly more often, whereas men were judged as high risk significantly more often. The three most frequently coded ‘other considerations’ differed for each gender. For men they were financial problems, lack of prospects for the future, and violent fantasies whereas for women they were forming a new intimate relationship, care for children, and prostitution.

Analyses of the predictive validity included two types of violence collapsed into a single outcome variable: (1) violent recidivism (operationalized with the HCR-20 definition of violence) after discharge was obtained from official judicial records for the “retrospective participants” and (2) data on inpatient violence was obtained from daily hospital information bulletins that detailed any disruptive incidents (incidents were coded only if they were acts of physical violence directed towards other persons). Values for all HCR-20 indices were higher for men than women. For men, AUC values for HCR-20 total and scale scores ranged from .75 to .88 and r s ranged from .42 to .62. For women,

AUCs ranged from .52 to .63 and r s ranged from .07 to .22. Values for final risk judgments were higher than values for the HCR-20 total and scale scores across both men (AUC = .91, r = .70) and women (AUC = .86, r = .57). Predictive indices for the PCL-R generally were lower than for the HCR-20.

Results indicate the predictive ability of the HCR-20 may be maximized when judgments of final risk are used rather than an actuarial approach wherein individual risk factors are summed.

Project and Scholarly Work

de Vogel, V. & de Ruiter, C. (2004). Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients. *The Journal of Forensic Psychiatry and Psychology*, 15, 145-164.

Summary

This research project assessed whether clinicians and researchers differ in their violence risk assessment of the same patients and whether raters' feelings towards the patients plays into their risk assessments. This study used the Dutch version of the HCR-20 on 60 patients (53 men and 7 women) in a Dutch psychiatric forensic hospital. The groups which coded the HCR-20 were comprised of 5 independent researchers, 7 treatment supervisor and 32 group leaders. The treatment supervisors were mostly clinical psychologist or psychotherapists. The group leaders were a diverse group with most having relevant higher vocational or academic training.

The mean HCR-20 scores were: Total score = 26.1 (SD = 6.5), H-scale = 14.6 (SD = 3.3), C-scale = 5.3 (SD = 2.2), R-scale = 6.1 (SD = 2.1). The inter-rater agreement was measured by *ICC*. The *ICC* between all three groups for the HCR-20 Total score was .79. For the H-scale the *ICC* was .82, the C-scale was .64, the R-scale was .57 and the final clinical risk judgment *ICC* was .65. Inter-rater agreements between subgroups of raters was equivalent to that of all three groups together.

In terms of differing scores by rater type, Group leaders rated significantly lower scores on the H-scale, Risk management items, and HCR-20 Total scores. There were no significant differences in the mean scores between the researchers and the treatment supervisors except for structured clinical risk judgments. Treatment supervisors more often judged patients as "low risk" compared to researchers.

Researchers stated that they spent about 120 minutes per risk assessment, group leaders spent about 30 minutes and supervisors about 15 minutes per assessment. Also, re-

searchers stated that they based their assessments predominantly on file information, whereas group leaders and treatment supervisors mostly relied on personal experiences with the patient.

Correlations between HCR-20 scores and a measure of feelings towards the patients showed many significant correlations. The HCR-20 total score was correlated with measures of patient's: helpfulness (r = -.28; p < .01), unhelpfulness (r = .38; p < .01), distant (r = .2; p < .05), accepting (r = -.19; p < .01), rejecting (r = .34; p < .01), and controlled (r = .46; p < .01). The HCR-20 risk judgment was correlated with measures of patient's: helpfulness (r = -.34; p < .01), unhelpfulness (r = .33; p < .01), close (r = .19; p < .5), distant (r = .4; p < .01), accepting (r = -.23; p < .01), rejecting (r = .34; p < .01), and controlled (r = .37; p < .01).

Stepwise multiple regression analyses showed that feelings of being controlled or manipulated by a patient significantly predicted high HCR-20 scores. 21% of the variance in the HCR-20 Total score was explained by feelings of being controlled by the patient. Also in stepwise regressions, feelings that the patient was close and distant predicted high risk judgments whereas feelings that the patient was helpful predicted low risk judgments. Together these three explained 23% of the variance in risk judgments.

Project and Scholarly Work

de Vogel, V., & de Ruiter, C. (2004). *Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20*. Manuscript under review.

Summary

This prospective study examined differences in accuracy between researchers (n = 9), treatment supervisors (n = 8), and group leaders (n = 59) with respect to individual versus consensus ratings and structured final risk versus actuarially based risk judgments. The sample comprised 127 men (a subset of whom were reported on previously; see de Vogel & de Ruiter, 2004) whose mean age at admission was 32.9 (SD = 9.6, range = 17-66). The index offenses were: 44% (attempted) homicide, 33% sexual offenses, 16% other violent offenses such as robbery, 7% arson. Mean length of stay in the hospital was 3.7 years (SD = 2.4, range = 0-12). More than half of the participants had abused substances in the past (8% alcohol, 15% drugs, and 44% multiple substances) and most had received previous psychiatric treatment.

Participants varied in terms of their treatment phase at the time the HCR-20 was coded. For participants who were commencing their first unsupervised leave from the hospital ($n = 9$), entering the transmural treatment phase ($n = 28$), or already were in the transmural treatment phase ($n = 24$), the R-scale was coded for the outside context. For participants who were newly admitted to the hospital ($n = 49$) and for existing inpatients ($n = 17$), the R-scale was coded for the context inside (risk of inpatient violence).

Raters coded the HCR-20 between January 2001 and June 2004 for each case independently and agreed upon a consensus score and a final risk judgment during a case conference. For 19 (15%) patients, more than one HCR-20 rating was completed because there was a change in their treatment phase. The most recent risk assessment was used for those participants.

Outcome data were obtained from daily information bulletins published in the hospital that report on inpatient violence and violence that occurred outside the hospital (e.g., for patients who were in the transmural treatment phase). The definition of physical violence was the same as that used in the HCR-20 manual. The mean follow up period was 21.5 months ($SD = 10.9$, range = 1-37). For individuals under mandated treatment conditions, data on violent recidivism was not obtained after the court order expired ($n = 20$; mean follow up period after discharge for this subgroup = 15 months, $SD = 8.8$, range = 4-34).

Group leaders gave significantly lower total and R-scale scores ($p < .05$) compared to researchers and treatment supervisors. There were no significant differences in mean HCR-20 scores between researchers and treatment supervisors. The mean HCR-20 consensus scores were higher (but not significantly so) than the mean HCR-20 scores of the three individual rater groups. Mean total scores were: researchers = 26.1 ($SD = 6.1$), treatment supervisors = 25.8 ($SD = 6.1$), group leaders = 24.1 ($SD = 5.8$), consensus = 26.8 ($SD = 5.6$). Mean H-scale scores were: researchers = 14.5 ($SD = 3.1$), treatment supervisors = 14.3 ($SD = 3.4$), group leaders = 14.0 ($SD = 3.4$), consensus = 14.8 ($SD = 3.1$). Mean C-scale scores were: researchers = 5.3 ($SD = 2.1$), treatment supervisors = 5.3 ($SD = 2.2$), group leaders = 5.0 ($SD = 2.0$), consensus = 5.5 ($SD = 2.1$). Mean R-scale scores were: researchers = 6.3 ($SD = 2.2$), treatment supervisors = 6.2 ($SD = 2.2$), group leaders = 5.3 ($SD = 2.2$), consensus = 6.4 ($SD = 1.9$).

There were no significant differences between the rater groups in final risk judgments. The percentages of low HCR-20 final risk judgments were: 24% researchers, 30% treatment supervisors, 21% group leaders, and 28% consensus. The percentages for judgments of moderate risk were: 45% researchers, 46% treatment supervisors, 43% group leaders, and 48% consensus. The percentages for

judgments of high risk were: 31% researchers, 24% treatment supervisors, 35% group leaders, and 24% consensus.

AUC values for physical violence for the total score were: researchers = .79 ($SD = .05$), treatment supervisors = .81 ($SD = .05$), group leaders = .75 ($SD = .05$), consensus = .85 ($SD = .04$). AUC values for the H-scale were: researchers = .73 ($SD = .06$), treatment supervisors = .74 ($SD = .06$), group leaders = .75 ($SD = .06$), consensus = .77 ($SD = .05$). AUC values for the C-scale were: researchers = .76 ($SD = .06$), treatment supervisors = .75 ($SD = .05$), group leaders = .66 ($SD = .06$), consensus = .80 ($SD = .05$). AUC values for the R-scale scores were: researchers = .74 ($SD = .06$), treatment supervisors = .71 ($SD = .05$), group leaders = .63 ($SD = .07$), consensus = .79 ($SD = .05$).

AUC values for the final risk judgment were: researchers = .77 ($SD = 2.2$), treatment supervisors = .75 ($SD = .05$), group leaders = .64 ($SD = .07$), consensus = .86 ($SD = .04$). Group leaders compared to researchers had a significantly lower AUC value for the final risk judgment ($\chi^2(1, N = 127) = 6.3, p < .01$). Group leaders' ratings compared to consensus ratings were significantly lower for the C-scale, R-scale, total score, and final risk judgment ($\chi^2(1, N = 127) =$ respectively 6.8, 4.9, 4.6 and 20.1, $p < .05$). The AUC value for the HCR-20 consensus final risk judgment was significantly higher than the individual final risk judgment of researchers, treatment supervisors and group leaders ($\chi^2(1, N = 127) =$ respectively 6.9, 5.3, and 20.1, $p < .01$).

Correlations for the HCR-20 total score were: researchers = .35, treatment supervisors = .36, group leaders = .30, consensus = .43. Correlations for the H-scale were: researchers = .27, treatment supervisors = .28, group leaders = .29, consensus = .32. Correlations for the C-scale were: researchers = .31, treatment supervisors = .31, group leaders = .19, consensus = .36. Correlations for the R-scale were: researchers = .29, treatment supervisors = .27, group leaders = .16, consensus = .35. Correlations for the final risk judgment were: researchers = .35, treatment supervisors = .33, group leaders = .19, consensus = .49. All p values $< .01$ for consensus, researchers, and treatment supervisors and at least $< .05$ for group leaders (except R-scale, $p = .16$).

Participants who scored above the median (27) relative to those below the median had significantly more had significantly more incidents of physical violence (Kaplan Meier log rank = 15.8, $p < .001$; odds ratio = 21.6, 95% CI = 2.8-167.2). Cox regression analyses with the three scales entered on the first block and final risk judgment entered on the second using the forward conditional method resulted in a significant model fit ($\chi^2(3, N = 127) = 22.9, p < .001$) at Block 1. HCR-20 final risk judgment demonstrated incremental validity as there was significant improvement to

the model's fit at Block 2 (χ^2 change (1, $N = 127$) = 6.8, $p < .01$).

AUC values and Pearson correlations were used to examine the predictive validity of consensus ratings for physical violence of the HCR-20 items. Items 2, 4, 5, and 7 from the H-scale, items 11, 12, 14, and 15 from the C-scale, and items 16, 17, and 19 from the R-scale had significant AUC values and correlations. Significant AUC values ranged from .67-.74 and significant correlations ranged from .21-.32. Cox regression analysis with all items included yielded a significant model (χ^2 (20, $N = 127$) = 43.7, $p < .01$). Using the forward conditional method to determine which HCR-20 items were significant predictors of incidents of physical violence produced a final model in which items 2 ($e^B = 6.4$, 95% CI = 1.5-28.0), 15 ($e^B = 3.4$, 95% CI = 1.5-8.1), and 17 ($e^B = 3.4$, 95% CI = 1.2-10.0) were significant predictors of incidents of physical violence.

The HCR-20 total score and final risk judgments were significantly predictive for both verbal abuse (total score: AUC = .72, SE = .05, $r = .36$, $p < .01$; final risk judgment: AUC = .65, SE = .05, $r = .28$, $p < .01$) and verbal threat (total score: AUC = .79, SE = .05, $r = .36$, $p < .01$; final risk judgment: AUC = .71, SE = .05, $r = .31$, $p < .01$).

Project and Scholarly Work

de Vogel, V., de Ruiter, C. Hildebrand, M., Bos, B. & van de Ven, P. (2004). Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health*, 3, 149-165.

Summary

The authors investigated the predictive validity, inter-rater reliability and survival rates while using the HCR-20 and PCL-R. The sample consisted of 120 patients discharged from a Dutch forensic psychiatric hospital between 1993 and 1999. The patients had a mean duration of treatment of 58.7 months and there was an average follow-up period of 73 months for this study. There were four different ways of discharge for these patients: transmural ($N = 30$; termination of treatment by court in line with hospital's advice and after a resocialization phase), conform advice ($N = 30$; termination of treatment by court in line with hospital's advice without resocialization phase), contrary to advice ($N = 30$; termination of treatment by court against the hospital's advice) and readmission to another institution ($N = 30$; readmission to another institution).

Inter-rater reliability was measured using ICC's. The ICC for the HCR-20 total score was .83. For the H-scale it was .89, for the C-scale it was .76, for the R-scale it was .58, and for the structured final risk judgment it was .73.

The mean scores for the HCR-20 and PCL-R by type of discharge were as follows. Transmural means were: PCL-R total (15.4), HCR-20 total (22.8), H-scale (12.6), C-scale (3.7), R-scale (6.5). Conform means were: PCL-R total (17), HCR-20 total (22.8), H-scale (12.8), C-scale (4.3), R-scale (5.6). Contrary means were: PCL-R total (20.2), HCR-20 total (27.6), H-scale (14.6), C-scale (5.4), R-scale (7.6). Readmission means were: PCL-R total (25.3), HCR-20 total (32), H-scale (16), C-scale (7), R-scale (9.1).

For the H-scale there were significant differences between the transmural and conform means as compared to the contrary mean ($p < .05$) and the contrary mean as compared to the readmission mean ($p < .05$). For the C-scale there were significant differences between the transmural and conform means as compared to the contrary and readmission means ($p < .05$). For the R-scale there were significant differences between the transmural and conform means as compared to the contrary and readmission means ($p < .05$). For the HCR-20 total score there were significant differences between the transmural and conform means as compared to the contrary and readmission means ($p < .05$). For the PCL-R total score, there were significant differences between the transmural and conform means as compared to the contrary mean ($p < .05$) and the contrary mean as compared to the readmission mean ($p < .05$).

Significant differences were found in the level of risk judgments given across the four discharge types. For the HCR-20, low risk judgments were given significantly more often to transmural and conform groups than to the readmission group ($p < .05$). Use of the HCR-20 also led to more moderate risk judgments for the transmural, conform and contrary groups as compared to the readmission group ($p < .05$). Lastly for the HCR-20, this measure led to more high risk judgments for the transmural and conform groups as compared to the contrary and readmission groups ($p < .05$). Using a cut-off of 26 on the PCL-R, there were higher judgments of risk given to those in the contrary and readmission groups as compared to the transmural or conform groups ($p < .05$).

Results showed that there were no significant differences between the transmural and conform or contrary groups in terms of violent recidivism. The conform group had a lower reconviction rate for violent offenses ($p < .05$), and the readmission group had a higher reconviction rate for violent offenses than the other three groups ($p < .01$).

The predictive validity of the HCR-20, PCL-R and clinical judgment for violent offending were calculated using AUCs and Pearson's correlations. AUC's: HCR-20 total score (.82; $p < .001$), H-scale (.80; $p < .001$), C-scale (.77; $p < .001$), R-scale (.79; $p < .001$), Risk judgment (.79; $p < .001$), PCL-R total score (.75; $p < .001$), PCL-R with cut-off of 26 or greater (.65; $p < .01$) and unstructured clinical judgment (.68; $p < .01$). Correlations: HCR-20 total score (.52; $p < .01$), H-scale (.47; $p < .01$), C-scale (.46; $p < .01$),

R-scale (.47; $p < .01$), Risk judgment (.51; $p < .01$), PCL-R total score (.43; $p < .01$), PCL-R with cut-off of 26 or greater (.39; $p < .01$) and unstructured clinical judgment (.32; $p < .01$).

The authors conclude that the HCR-20 structured final judgment was significantly more accurate than unstructured clinical judgment in predicting violent recidivism ($p < .05$). The HCR-20 was also significantly more accurate than the PCL-R in predicting violent recidivism ($p < .05$) except when the item H7 (psychopathy) was removed from the HCR-20 total score ($p = .08$).

Project and Scholarly Work

de Vogel, V., de Ruiter, C., & Vandeputte, C. (2001, November). *Implementation of the HCR-20 and SVR-20 in a Dutch forensic psychiatric hospital*. Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden.

Summary

This research project assessed the reliability and predictive validity of the HCR-20 and the SVR-20 (Sexual Violence Risk-20). The project also assessed who would be the most suitable to perform risk assessments. The study used 60 patients (53 males, 7 females), assessed them initially before their entrance into a transmurial phase and then again in the transmurial phase.

The mean HCR-20 scores were: Total score = 26.1 ($SD = 6.5$), H-scale = 14.6 ($SD = 3.3$), C-scale = 5.3 ($SD = 2.2$), R-scale = 6.1 ($SD = 2.1$). Inter-rater reliability was assessed using *ICCs*. Across assessors, treatment leaders and group leaders together, the *ICCs* were as follows: HCR-20 total score (.79), H-scale (.82), C-scale (.64), R-scale (.57), and final structured risk judgment (.65). In terms of differing scores by rater type, those who were assessors gave the highest HCR-20 scores, with treatment supervisors giving the next highest and group leaders giving the lowest. Significant differences only existed, though, between assessors and group leaders.

Inpatients' H, C, R and Total scores were higher than those in the transmurial phase. Inpatient final risk judgments were higher than when in the transmurial phase.

Project and Scholarly Work

Fujii, D., Lichten, A., & Tokioka, A. (2004). *Structured professional judgment versus actuarial data in violence risk prediction using the Historical Clinical Risk Management-20*. Manuscript under review.

Summary

The accuracy of actuarial predictions of inpatient violence using a cut-off score of 27 on the HCR-20 versus clinicians' structured professional judgments (SPJ) was compared. The HCR-20 was administered by three doctoral level psychologists within the first week of arrival to 169 patients (138 men and 31 women) admitted consecutively to a state hospital between February 2002 and January 2003. The most common admission diagnoses were schizoaffective (18%) and paranoid schizophrenia (16%). The inter-rater reliability coefficient for 12 cases was .94. Episodes of inpatient violence (operationalized by the definition of violence in the HCR-20 manual) were recorded from hospital event records for a minimum of three months post-admission.

For actuarially derived predictions, the hit rate = 71%, sensitivity = 30% (12/40), specificity = 86% (95/111), positive predictive power (PPP) = 43% (12/28), negative predictive power (NPP) = 77% (95/123), and AUC = .61 (range: .51-.72). SPJ-based predictions (patients were rated either as high risk or low/moderate risk) generally were higher: hit rate = 77%, sensitivity = 45% (18/40), specificity = 88% (105/120), PPP = 55% (18/33), NPP = 83% (105/127), and AUC = .70 (range: .56-.77).

A step-wise regression was completed using the number of violent inpatient episodes as the criterion variable and overall HCR-20 scores and five-level SPJ predictions (low, low-moderate, moderate, moderate-high, and high) as the predictor variables. SPJ-based predictions added incremental validity over actuarial predictions (an increase in r^2 from .036 to .092, $p < .05$), whereas the reverse was not true.

The study also reports on clinicians' predictions regarding the situational contexts in which violence might occur for each participant based on his or her historical background. Results provide support for the use of the SPJ approach in making predictions of inpatient violence among forensic psychiatric patients.

Project and Scholarly Work

Fujii, D., Tokioka, A., & Lichten, A. (in press). *Cultural differences in violence risk prediction of psychiatric inpatients using the Historical Clinical Risk Management-20*. *Psychiatric Services*.

Summary

The authors noted that the majority of studies on the HCR-20 have used samples that comprise predominantly Caucasians of European heritage. The purpose of the present study was to examine retrospectively cultural differences in violence risk assessment of psychiatric inpatients using

the HCR-20. Participants were drawn from a sample of 169 consecutive admissions (the same sample reported on by Fujii, Lichten, & Tokioka, under review). Participants were included in this study if they described themselves as Asian-American (AA; $n = 51$), Euro-American (EA; $n = 46$), or Native American of part-Hawaiian (NAH, $n = 38$) heritage. Participants were considered AA if their ethnicity was Japanese, Chinese, Korean, Filipino, or Vietnamese and NAH if they reported Hawaiian as one of their ethnic languages. Participants with a mixed ethnic heritage, apart from the NAH group, were excluded. The final sample consisted of 88 men and 20 women and had a mean age of 40.1 years ($SD = 12.6$) and a mean education level of 11.9 years ($SD = 2.5$).

There were no differences in rates of institutional violence (i.e., threats or assaults on patients and staff) among the three ethnic groups. ROC analyses indicated the highest accuracy for predicting inpatient violence was obtained for the NAH group ($AUC = .730$) and the lowest accuracy for the AA group ($AUC = .575$; AUC for the EA group = $.638$). Stepwise multiple regressions were conducted for each ethnic group using HCR-20 items as predictor variables and the number of violent events (multiplied by \log_{10} to control for a skewed distribution) as the criterion variable. Results indicated a unique pattern of predictors was associated with each cultural group. Models for AA and EA each produced a single significant predictor. For AA, item C4 (impulsivity) accounted for 16.1% of the variance. For EA, item H2 (young age at first violence) accounted for 13.3% of the variance. The largest effect size ($R^2 = .430$) was obtained for NAH, which had three significant predictors (H2, young age at first violence; H3, relationship instability; and R1, plans lack feasibility).

Results are discussed in terms of possible explanations for the disparities in observed predictive ability of the HCR-20 as a function of ethnicity.

Project and Scholarly Work

Grann, M., Belfrage, H., & Tengström, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behavior*, 27, 97-114.

Summary

This was a retrospective follow-up of 404 forensic patients who had committed violent offences in Sweden, and who were followed up for a period of two years. This study compared to predictive characteristics of the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993) and the H Scale of the HCR-20. The sample was further broken down into two sub-samples or cohorts: 1) 293 violent offenders with ICD-9 diagnoses of personality disorder;

2) 111 violent offenders with diagnoses of schizophrenia.

Across both groups, the AUC of the ROC for the H Scale was $.71$ ($95\% CI = .66 - .76$). At the cut-off score of 12 on the H Scale (the inflexion point), sensitivity = $.71$; specificity = $.61$; positive predictive power = $.35$, and negative predictive power = $.88$. For the VRAG, the AUC was $.68$ ($95\% CI = .63 - .73$). At the cut-off score of 13 on the VRAG (the inflexion point), sensitivity = $.50$; specificity = $.77$; positive predictive power = $.39$, and negative predictive power = $.84$.

In the personality disordered cohort, the AUC of the ROC for the H Scale was $.71$ ($95\% CI = .66 - .76$). At the cut-off score of 12 on the H Scale (the inflexion point), sensitivity = $.72$; specificity = $.60$; positive predictive power = $.38$, and negative predictive power = $.86$. For the VRAG, the AUC was $.68$ ($95\% CI = .62 - .73$). At the cut-off score of 13 on the VRAG (the inflexion point), sensitivity = $.57$; specificity = $.71$; positive predictive power = $.40$, and negative predictive power = $.83$.

In the schizophrenia cohort, the AUC of the ROC for the H Scale was $.66$ ($95\% CI = .56 - .75$). At the cut-off score of 8 on the H Scale (the inflexion point), sensitivity = $.88$; specificity = $.36$; positive predictive power = $.19$, and negative predictive power = $.95$. For the VRAG, the AUC was $.60$ ($95\% CI = .50 - .69$). At the cut-off score of 0 on the VRAG (the inflexion point), sensitivity = $.68$; specificity = $.53$; positive predictive power = $.20$, and negative predictive power = $.91$.

Grann et al. concluded that both the H Scale and the VRAG predicted violence significantly better than chance (except for the VRAG in the schizophrenia group). They comment that the obtained values could under-represent the actual predictive accuracy of the instruments because several items on each scale had to be "approximated." The sensitivity of the H Scale tended to be greater than that for the VRAG, whereas the specificity of the VRAG tended to be greater. Among the schizophrenia group, only the H Scale was better than chance.

Project and Scholarly Work

Grann, M., & Långström, N. (in press). Actuarial assessment of risk for violence: To weigh or not to weigh? *Criminal Justice and Behavior*.

Summary

Using data from Grann et al. (2000), this investigation evaluated the relative accuracies of different options for weighting H scale scores. The authors used five approaches: nonweighted, Nuffield approach, logistic regression model (one-by-one), logistic regression model (11-

term algorithm), and artificial neural network. They split the sample into training (or calibration) and validation seeds or subsets. Results showed that the unweighted procedure produced the largest average AUC value (.72), compared to the Nuffield approach (.71), logistic regression one-by-one (.71), logistic regression 11-term algorithm (.68) and artificial neural network (.64). These findings are consistent with research showing that unweighted predictors are often as accurate as optimally-weighted procedures.

Project and Scholarly Work

Gray, N. S., Snowden, R. J., MacCulloch, S., Phillips, H., Taylor, J. & MacCulloch, M. J. (2004). Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally disordered offenders: A comparative study of HCR-20, PCL:SV, and OGRS. *Journal of Consulting and Clinical Psychology, 72*, 523-530.

Summary

This study compared the predictive accuracy of the HCR-20, PCL:SV, and the Offender Group Reconviction Scale (OGRS; Copas & Marshall, 1998; this is a criminogenic risk assessment tool based on six demographic and offending history variables that estimates the probability of reconviction within 2 years of release) among 315 forensic psychiatric patients discharged from a medium-secure facility in South Wales, United Kingdom between 1992 and 1999. The sample primarily was male (87.6%), Caucasian (84.4%; 12.4% Black Caribbean or Black African; 1.3% Asian; .6% 'mixed'; 1.3% 'unknown'), and diagnosed with schizophrenia or psychotic disorder (49.2%; 16.8% personality disorder; 9.8% affective disorder; 6.3% drug induced psychosis; 5.1% MR; 1.0% substance misuse disorder; 3.2% 'other'; 8.6% 'unknown').

Two psychologists blind to outcome completed all assessments using file information available at discharge, which consisted of mental health, criminal, social work, and probation records. Participants were followed up for at least two years (mean = 6.00 years, *SD* = 1.77 years). During the follow-up period, 36.5% were convicted of any type of offense.

Mean HCR-20 scores were: Total (19.90, *SD* = 7.02; range 0-36); H-scale (11.39, *SD* = 3.97; range 0-20); C-scale (3.77, *SD* = 2.42; range 0-10); R-scale (4.68, *SD* = 2.63; range 0-10). Mean PCL:SV scores were: Total (8.25, *SD* = 5.18; range 0-20), Part 1 (3.79, *SD* = 3.79, range 0-11), and Part 2 (4.50, *SD* = 2.83; range 0-12). The mean OGRS score was .49 (*SD* = .29; range .03 - .99). HCR-20 and PCL:SV total and scale/Part scores correlated highly and significantly with one another (ranging from .36 to .78).

Correlations for the OGRS with the HCR-20 and PCL:SV tended to be lower and were not consistently significant.

Survival analysis revealed that 87% of the offenses occurred within approximately 3 years. The Mantel-Cox log-rank statistic was used to evaluate the percentage of patients in low, medium, and high risk predictor groups who committed an offense following discharge for the three measures. For the PCL:SV, risk groups were defined as follows: low (scores of 12 or less); medium (scores of 13-17); and high (scores of 18 or more). The distribution of scores was used to trisect the sample into groups for the other two measures. For the HCR-20, groups were defined as follows: low (scores of 16 or less); medium (scores of 17-22); and high (scores of 23 or more). Groups for the OGRS were: low (<.29); medium (.29 - .67); and high (>.67). Although significant results were obtained with respect to any type of offending outcome for all measures, the log-rank value for the OGRS was much higher (83.78) than the values for the HCR-20 (10.70) and PCL:SV (10.76).

Mean scores on the three measures were compared across participants who offended and those who did not. Cohen's *d* values were as follows: HCR-20 total (.35), H scale (.38), C scale (-.08), R scale (.41), PCL:SV total score (.54), Part 1 (.25), Part 2 (.70), and OGRS (1.28).

Using ROC analysis, total scores of all three measures were associated significantly with offending outcome (AUC values for the HCR-20, PCL:SV, and OGRS were .61, .66, and .81, respectively). AUC values for the subscales were more variable, with the H scale (.62), R scale (.62), and PCL:SV Part 2 (.72) reaching significance, but with the C scale (.48) and PCL:SV Part 1 (.57) failing to do so. ROC analyses that examined serious and minor offenses revealed a similar pattern of results. When participants were divided into groups on the basis of diagnosis (i.e., mental illnesses, personality disorders, and 'other' diagnoses that included mental retardation, developmental disorder, and physical diagnoses), the size of the AUC values for the mental illness and 'other' groups was similar to the above-described values for the overall sample (although none of the values except for the OGRS were significant for the 'other' group).

Finally, a logistic regression analysis was undertaken to investigate whether the HCR-20 and/or PCL:SV could make an additional significant contribution to an OGRS-only model. Using a forced-entry method, no total or scale/Part variables added incremental validity.

The discussion section reiterated the findings and noted that the timing at which the C scale was scored (i.e., prior to discharge when symptomatology was as low as it likely ever would be, rather than during a time of active symptomatology) may have impacted the findings. The authors concluded that adoption of a singular focus on mental

health factors ignores important sources of information predictive of reoffending.

Project and Scholarly Work

Gretenkord, L., Muller-Isberner, R., Ozokuy, K. & Sommer, J. (2002, March). *Validating the HCR-20: relationship between levels of security and the CR-10 score in hospital order treatment*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.

Summary

This study used a prospective design to determine the relationship between the HCR-20 and levels of security in a forensic psychiatric hospital. Over 12 months, they followed 220 individuals (209 men and 11 women) who had a hospital order sentence and had been hospitalized for the entire 12 month period. They predicted that if risk factors were changed by successful treatment, the dynamic part of the HCR-20 (the CR-10) should decrease.

The following significant correlations between individual C and R scale items and level of security were found over time: C1 ($r = -.286$), C2 ($r = -.264$), C4 ($r = -.236$), C5 ($r = -.347$), R1 ($r = -.42$), R2 ($r = -.443$), R3 ($r = -.237$), R4 ($r = -.409$), R5 ($r = -.227$). The C-scale in total showed a significant correlation with level of security over time ($r = .369$; $p < .001$), as did the total R-scale ($r = .575$; $p < .001$) and to a lesser extent, the H-scale ($r = .167$; $p < .05$). The PCL also showed a significant correlation with level of security over time ($r = .227$; $p < .01$).

The authors conclude that there are robust correlations between CR-10 items and levels of security and that the CR-10 seems to be a good indicator of treatment progress.

Project and Scholarly Work

Grevatt, M., Thomas-Peter, B., & Hughes, G. (2004). Violence, mental disorder and risk assessment: Can structured clinical assessments predict the short-term risk of inpatient violence? *The Journal of Forensic Psychiatry & Psychology*, 15, 278-292.

Summary

This study examined retrospectively the predictive ability of the combined HCR-20 H- and C-scales and the Violence Risk Scale 2 (VRS; Wong & Gordon, 2001) within the first six months of admission to a forensic unit. The VRS comprises six static and 20 dynamic factors rated on a 0 (not present/not applicable) to 3 (definitely present/applicable). The measures were completed retrospec-

tively for 44 men using information available at admission. One rater, who was blind to outcome of institutional violence, completed the HC composite and VRS. Another rater, who was blind to risk assessment ratings, rated the incidents of violence. Types of violence coded were physical assault, verbal aggression, and damage to property.

Mean scores on the HC composite were: full scale (19.44, $SD = 3.45$); H-scale (13.15, $SD = 3.25$); and C-scale (6.05, $SD = 1.98$). Total scores on the HC composite and VRS (prorated for omitted items) did not distinguish participants who were aggressive in the institution from those who were not nonaggressive. ROC analyses indicated that the HC and VRS indices, with the exception of the C-scale, tended to not have predictive accuracy for inpatient violence that was greater than chance (the highest value was for the HC composite for physical assaults, $AUC = .56$, $SD = .10$). AUC values for the C-scale were larger: any incidents (.72, $SD = .08$); physical assaults (.60, $SD = .11$); verbal abuse (.81, $SD = .07$); and damage to property (.65, $SD = .10$).

Four multiple regression analyses were conducted (one for each category of violence as the dependent variable) using the measures' subscales (i.e., H-scale, C-scale, VRS static, and VRS dynamic) as the predictors. C-scale was the only significant predictor for any institutional incidents and was the most significant predictor for verbal assault. None of the subscales emerged as significant predictors for the outcomes of physical assault and damage to property.

When individual items that comprise the HC composite and VRS scale were considered, those most predictive of inpatient violence were HC composite items that assess a previous diagnosis of mental illness, lack of insight, and active signs of mental illness. Protective factors for institutional violence included VRS items that assess relationship instability, number of young offender convictions, violent lifestyle, and violence throughout the lifespan.

Project and Scholarly Work

Hilterman, E. & Chakhssi, F. (2002, March). *Prospective assessment of risk: comparing HCR-20, Behavioural Status Index (BSI) and Leave Risk Assessment (LRA)*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.

Summary

This study was a prospective analysis comparing the HCR-20 (Dutch version), BSI and LRA in their ability to assess future risk. The HCR-20 was given before the first supervised leave request and before every extension of unsuper-

vised leave. The BSI was given every half year before treatment evaluation. The LRA was given before extension leave trajectory and advice regarding the extension of the hospital order.

Inter-rater reliability for the HCR-20 ($N = 11$) was: H-scale ($ICC = .92$), C-scale ($ICC = .91$), R-scale ($ICC = .95$) and Total score ($ICC = .98$). Inter-rater reliability for the BSI ($N = 75$) was: Direct aggression ($ICC = .84$), Obstructionism ($ICC = .84$) and BSI Risk ($ICC = .89$). Inter-rater reliability for the LRA ($N = 14$) was: LRA-SV ($ICC = .99$) and LRA-DV ($ICC = .84$).

The distribution of scores for the measures was as follows. With a sample size of 27, the HCR-20 had a mean of 25.2 ($SD = 7.54$). With a sample size of 62, the BSI had a mean of 4.57 ($SD = .37$). With a sample size of 16, the LRA-SV had a mean of .6 ($SD = 1.72$) and the LRA-DV had a mean of -5 ($SD = 3.34$).

Correlations between the HCR-20 and the BSI were conducted. The HCR-20 total score was correlated with BSI-Obstructionism ($r = -.38$; $p < .1$), BSI-Direct aggression ($r = -.37$; $p < .1$) and with BSI-Risk ($r = -.43$; $p < .05$). The HCR-20 H-scale was correlated with BSI-Direct aggression ($r = -.35$; $p < .1$) and with BSI-Risk ($r = -.36$; $p < .1$). The HCR-20 C-scale was correlated with BSI-Obstructionism ($r = -.51$; $p < .01$), BSI-Direct aggression ($r = -.46$; $p < .05$) and with BSI-Risk ($r = -.56$; $p < .01$). The HCR-20 R-scale was not correlated with any BSI score. The HCR-20 total score was correlated with LRA-SV scale ($r = .77$; $p < .01$) and with the LRA-DV scale ($r = .5$; $p < .1$). The HCR-20 H-scale was correlated with LRA-SV scale ($r = .73$; $p < .01$) but not with the LRA-DV scale. The HCR-20 C-scale score was correlated with LRA-SV scale ($r = .7$; $p < .01$) and with the LRA-DV scale ($r = .52$; $p < .1$). The HCR-20 R-scale was correlated with LRA-SV scale ($r = .58$; $p < .05$) but not with the LRA-DV scale.

Project and Scholarly Work

Müller-Isberner, R., & Jockel, D. (1997, September). *The implementation of the HCR-20 in a German hospital order institution*. Paper presented at the Seventh European Conference on Psychology and Law, Solna, Sweden.

Summary

100 forensic psychiatric patients were rated on the German version of the HCR-20 (which includes 3 variables not in the original version). There were 96 men, and the mean age of the sample was 38.8 years. Only the H and C scales were rated. Most index offences were of a violent nature: homicide (24%); severe bodily harm (21%); violent sexual offences (20%); arson (13%); and 24 other offences. Close

to half (43%) of the sample had primary diagnoses of functional psychosis.

Two psychiatrists rated a subsample of 45 offenders, allowing interrater reliability analyses. For the H Scale items, Kappa ranged from .54 to 1.00, with a mean Kappa of .89. In 91% of cases, the two clinicians were within one point on ratings of H Scale total scores. Kappa was not as high for the C Scale, ranging from .33 to .65, with a mean Kappa of .49. In 71% of cases, clinicians were within one point on the C Scale.

Mean H scores were greatest for personality disordered patients with low IQs ($M = 13.6$) and lowest for patients with major brain damage ($M = 9.5$). Homicide offenders ($M = 9.5$) and nonviolent sexual offenders ($M = 8.0$) scored lowest on the H Scale, whereas patients who had committed "violent property offences" scored highest ($M = 13.8$). There were no differences on the C Scale as a function of index offence.

Project and Scholarly Work

Müller-Isberner, R., Sommer, J., Özokuy, K., & Freese, R. (1999, November). *Clinical use of the HCR-20 for predicting violence in a German Forensic Psychiatric Hospital*. Paper presented at the International Conference on Risk Assessment and Management: Implications for Prevention of Violence, Vancouver, BC, Canada.

Summary

Coded the German Version of the HCR-20 on 220 forensic psychiatric patients (209 male). Patients had committed serious offences, been found not criminally responsible, and had been judged to have a high risk for recidivism. Hospitalization is indeterminate; court requires annual progress reports. Mean age of sample was 38.1 ($SD = 10.1$). Index offences were as follows: homicide (24% of males; 18% of females); assault (21% of males; 27% of females); sexual offences (29% of males; 0% of females); arson (9% of males; 55% of females); property and other offences (15% of males; 0% of females). Diagnostic categories for males were 45% major mental disorder, 35% personality disorder, 20% brain damage, mental retardation or substance abuse disorders. For females, diagnostic categories were 55% major mental disorder, 18% personality disorder, 27% mental retardation. Mean (SD) scores: Total (24.87; 5.90); H (11.97; 3.42); C (5.30; 2.18); R (7.58; 1.86).

Researchers carried out interrater reliability data by having 7 "experienced psychiatrists" rate 50 patients. Cohen's Kappa for chance-corrected agreement on categorical final risk judgments was .72.

Numerous correlations between H, C, R, PCL:SV, and various inpatient indices of aggression were reported separately for patients with primary diagnoses of major mental disorder versus personality disorder. Correlations between predictors and outcome for patients with major mental disorders were as follows: Minor aggressive acts: threats (HCR-20 Total = .39; H, C, & R = .22, .44, .30; PCL:SV = .30); insults (HCR-20 Total = .30; H, C, & R = ns, .36, .21; PCL:SV = .28). Medium aggressive acts: willful property damage (HCR-20 Total = .40; H, C, & R = .23, .51, .27; PCL:SV = .24); terror/incitement (HCR-20 Total = .20; H, C, & R = ns, .27, ns; PCL:SV = .21). Major aggressive acts: physical violence toward staff (HCR-20 Total = .23; H, C, & R = ns, .34, ns; PCL:SV = ns); sex offences (HCR-20 Total = .20; H, C, & R = ns, .25, ns; PCL:SV = .21). No measure correlated with firesetting or physical violence toward patients. Correlations for the personality disordered patients were similar for minor aggressive acts, and less consistent for other outcomes.

The researchers concluded that both the HCR-20 and PCL:SV did not predict serious violence consistently. The C-Scale was most consistent for patients with major mental disorders; the PCL:SV for patients with personality disorders alone. Possible reasons include low base rates or small N (neither were reported). The authors claimed that the accuracy of measures for serious violence might have been affected by staff taking measures to prevent violence (hence reducing base rates and likely affecting the behaviour of patients). Staff may have prevented the violence of higher risk patients, hence reducing the correlations between high scores and high incidents of violence.

Project and Scholarly Work

Nicholls, T. L., Vincent, G. M., Whittemore, K. E., & Ogloff, J. R. P. (1999, November). *Assessing risk of inpatient violence in a sample of forensic psychiatric patients: Comparing the PCL:SV, HCR-20, and VRAG*. Paper presented at the International Conference on Risk Assessment and Risk Management, Vancouver, British Columbia, Canada.

Summary

This study is based on the same data set as Vincent (1998), *infra*, but addressed independent issues. The reader is referred to the annotation of Vincent (1998) for a description of the general methodological factors.

This research assessed the predictive ability of the HCR-20, VRAG, and PCL:SV in terms of inpatient violence of 125 forensic psychiatric patients. The authors carried separate analyses for pre-disposition and post-disposition time periods (i.e., pre- and post Review Board hearing). Violence was categorized as verbal, physical, and “any,” and

was coded from detailed files. Analyses included univariate Pearson r correlations, ROC analyses, and hierarchical logistic regression analyses.

Pre-disposition violence. For the HCR-20, Pearson r values for verbal, physical, and any violence were as follows: .39, .36, .46. These generally were higher than for the VRAG (.22, .07, .21) or the PCL:SV (.25, .26, .32). The AUCs for the HCR-20 were .72, .72, and .77, and again were generally higher than for the VRAG (.62, .66, .69) or PCL:SV (.65, .54, .62). Hierarchical logistic regression showed that the PCL:SV predicted any and physical inpatient violence when entered as alone in Block 1, the VRAG did not add to this on Block 2, and, on Block 3, only the HCR-20 predicted violence (the PCL:SV was no longer significant, nor was the VRAG). Results were not reported for verbal violence.

Post-disposition violence. For the HCR-20, Pearson r values for verbal, physical, and any violence were as follows: .31, .31, .36. These generally were higher than for the VRAG (.20, .08, .23) or the PCL:SV (.20, .14, .16). The AUCs for the HCR-20 were .68, .69, and .71, and again were generally higher than for the VRAG (.62, .55, .63) or PCL:SV (.60, .58, .59). Hierarchical logistic regression showed only the HCR-20 predicted any and physical violence (the PCL:SV and VRAG were not significant in any Block). Results were not reported for verbal violence.

Project and Scholarly Work

Pham, T. (2001, November). *Assessing risk for violence among Belgian offenders*. Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden.

Summary

This research was a mixed time perspective study using the HCR-20, PCL-R and the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993). This study used 80 subjects to measure correlations of the above measures with type of offense, and a subset of 58 subjects to conduct Kaplan-Meier survival analyses and ROCs. Type of offense fall under two categories: general recidivism (any offense committed after release) and violent recidivism (homicide, assault and battery, theft with violence, any sex offense). The mean follow up period after release was 994 days.

In terms of general recidivism, the PCL-R was correlated ($r = .26$; $p < .05$) with drug offenses, ($r = .33$; $p < .01$) with carrying a weapon, and ($r = .46$; $p < .01$) with theft. The HCR-20 was correlated ($r = .24$; $p < .05$) with drug offenses, ($r = .23$; $p < .05$) with carrying a weapon, and (r

=.40; $p < .01$) with theft. The VRAG was correlated ($r = .26$; $p < .05$) with drug offenses and ($r = .47$; $p < .01$) with theft.

In terms of violent recidivism, the PCL-R was correlated ($r = .39$; $p < .01$) with assault and battery and ($r = .48$; $p < .01$) with violent theft. The HCR-20 was only correlated ($r = .32$; $p < .01$) with violent theft. The VRAG was correlated ($r = .29$; $p < .05$) with assault and battery and ($r = .38$; $p < .01$) with violent theft.

In terms of predicting general recidivism, the PCL-R had an *AUC* of .78. The VRAG had an *AUC* of .86 and the HCR-20 had an *AUC* of .79. With predicting violent recidivism, the PCL-R had an *AUC* of .85, the VRAG had an *AUC* of .84, and the HCR-20 had an *AUC* of .78.

The following Pearson correlations between the measures were found: PCL-R was correlated with the VRAG ($r = .67$) and the HCR-20 ($r = .83$), while the VRAG was correlated with the HCR-20 ($r = .68$).

Project and Scholarly Work

Philipse, M. (2002, March). *Post-dictive validity of the HCR-20 in a Dutch forensic psychiatric sample*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.

Summary

This was a retrospective validation study in the Netherlands of the Dutch version of the HCR-20 (Philipse, de Ruiter, Hildebrande & Bauman, 2000). This research study used subset of 69 patients from three hospitals from a larger prospective study on assessing risk for re-offending. The research was conducted without using Item 7 (Psychopathy) from the H scale. The sample consisted of 64 males and 5 females. The types of offenses were categorized as violent, sex, and arson. Patients had left the hospital between 1/1/96 and 12/31/98. Re-offending data was collected on 1/22/02 with an average of 4 years and 4 months of time for patients to have been outside of the hospital. 21 (30%) had been found to have had renewed contact with the law.

Inter-rater reliability for the HCR-20 (Dutch version) was: *ICC* HCR-20 Total = .90, *ICC* H-scale = .79, *ICC* C-scale = .76, *ICC* R-scale = .67. The total and R-scale scores were significantly lower for patients discharged from the hospital according to hospital advice. R-scores were predictive of type of discharge (*AUC* = .67). HCR-20 (Dutch version) was most effective for non-sexual offenders. Deleting females did not alter the findings. The postdictive validity *AUCs* for committing a violent act (when excluding sex offenders) were: HCR-20 total score = .67, H-scale = .72,

C-scale = .60, R-scale = .58. The postdictive validity *AUC* for clinical judgment was .64, as was the number of previous convictions. Reducing the HCR-20 into smaller units increased the postdictive *AUC* values. Using only the H2, H4, H5, H10, C3 and C4 items achieved an *AUC* of .82. Using only the H2, H5, H10 and the C4 items achieved an *AUC* of .90.

In terms of decision making, the 4-item version of the HCR-20 with a cut-off of 50% identified all offenders with 2.2 false positives per true positives. The 4-item version of the HCR-20 with a cut-off of 80% identified 5 of 8 offenders with .6 false positives per true positives. Implications for the clinical assessment of risk of re-offending and the best composition of the HCR-20 items are discussed.

Philipse, M., Erven, T. van, & Peters, J. (2002). *Risicotaxatie in de tbs: van geloof naar empirie*. [Risk assessment in tbs: from belief to empiricism.] *Justitiële Verkenningen [Judicial Explorations]*, 28(8), 77-93.

Summary

Risk assessment in Dutch forensic psychiatry (tbs) is still dominated by an unstructured clinical approach. Researchers have argued in favour of a standardised approach because international research reports limited predictive validity of clinical approaches. The Dutch version of the clinical-actuarial debate is briefly summarised in this article. A study is presented that evaluates the validity of an international risk assessment tool, the HCR-20 in tbs. This shows that using the HCR-20 may improve risk assessment under certain conditions, although unstructured clinical judgement performs quite well too. Also, it is shown that clinically adjusted HCR-scores are slightly better than actuarial scores. However, in the final analysis historical predictors outperform all other measures. It is concluded that the HCR-20 may constitute a meaningful addition to Dutch risk assessment practice, though it is imperative that all persons dealing with this and similar instruments have a clear view of their limitations.

Project and Scholarly Work

Scharin, C. (1999). *Bedömning av återfallsrisk hos rättspsykiatriskt undersökta personer: En utvärdering av skattningsskalan HCR-20*. Unpublished manuscript.

Summary

The Swedish version of the HCR-20 was coded on 49 forensic psychiatric patients. [Sample characteristics unavailable at this time until English translation available]. Proportion of violence in various score categories was calculated for the total HCR-20 score and the H scale alone.

Results were as follows: HCR-20 total score from 0-19, 15% violent; total score from 20 to 40, 64% violent. H scale score of 0 to 5 (0% violent), 6 to 10 (31% violent), 11-15 (54% violent), 16 to 20 (80% violent).

Project and Scholarly Work

Strand, S., & Belfrage, H. (2001). Comparison of HCR-20 scores in violent mentally disordered men and women: Gender differences and similarities. *Psychology, Crime and Law, 7*, 71-79.

Summary

The purpose of this study was to compare the scores on the HCR-20 between male and female forensic patients. Using the official Swedish translation of the HCR-20, all female patients ($n = 63$) who entered a Swedish forensic facility over 10 years were assessed with file, and, where possible, also with interview. Comparisons were made with all 85 male patients admitted to two Swedish forensic hospitals in 1998.

The female sample was younger (30.8 vs. 35.1 years), more often diagnosed with a personality disorder (55.6% vs. 36.5%, specifically borderline [85.7% vs. 25.8%], and less often antisocial [0.0% vs. 25.8%]). Females were less often admitted after committing violent crimes (9.5% vs. 31.8% murder; 17.5% vs. 31.8% other violent crimes), and more often admitted from general psychiatry due to violence (42.9% vs. 2.4%).

There were no differences in scale or total scores between genders. Total score = 24.76 ($SD = 6.95$) female, 25.51 ($SD = 7.92$) male; H scale = 12.94 ($SD = 3.58$) female, 13.81 ($SD = 4.21$) male; C scale = 5.11 ($SD = 2.57$) female, 5.00 ($SD = 2.48$) male, R scale = 6.71 ($SD = 2.85$) female, 6.68 ($SD = 2.80$) male.

There were differences on some of the items, likely reflecting the general differences between genders. Males scored higher on Previous Violence (H1), Young Age... (H2), Substance Use Problems (H5), and Negative Attitudes (C2). Females scored higher on Personality Disorder (H9), Impulsivity (C4), and Stress (R5).

Project and Scholarly Work

Strand, S., Belfrage, H., Fransson, G., & Levander, S. (1999). Clinical and risk management factors in risk prediction of mentally disordered offenders: More important than actuarial data? *Legal and Criminological Psychology, 4*, 67-76.

Summary

The Swedish version of the HCR-20 was coded on 40 male forensic psychiatric patients in a postdictive study of the HCR-20 and PCL:SV. There were 22 recidivistic patients and 18 non-recidivistic patients who were matched on demographic, clinical, and criminal variables. The rater was blind to recidivism status. Overall, the recidivistic group scored 8 points higher than the nonrecidivistic group (M s and SD s = 30.77 [7.22]; 22.39 [6.85], respectively). Although not reported, this represents a Cohen's d of 1.19, which is a large effect size. All persons ($n = 11/40$) with scores above 34 on the HCR-20 recidivated. Interestingly, Strand et al. report that for patients who scored between 24 and 28, prediction was random. However, all recidivistic patients in this scoring range scored 2/2 on R5 (Stress), and this item alone differentiated the two groups (for this scoring range).

The area under the curve of the receiver operating characteristic analysis was .80 for the HCR-20, and .70 for the PCL:SV. Using a cut-off score of 29/40 on the HCR-20, sensitivity was reported to be .89 and specificity .64. With a cut-off of 17/24 on the PCL:SV, sensitivity was .89, and specificity was .59.

Surprisingly perhaps, the items from the Clinical and Risk Management scales were much stronger in separating the two groups than was the Historical scale. Strand et al. point out that this finding may stem from the fact that the patients in their sample, given their offences and dispositions to a forensic hospital, were homogenous on historical factors.

Project and Scholarly Work

Tengström, A. (2001). Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry, 55*, 243-249.

Summary

This was a long-term predictive validity assessment (with retrospective data collection) of the VRAG and the Historical part of the HCR-20. The sample consisted of 106 violent offenders with schizophrenia in Sweden. The mean detention time for the offenders was 18 months. All subjects were followed from discharge or start of probation until each subject was at least 5 year from that point. The average time from discharge at follow-up was 86 months. The definition used for this study of violent recidivism was closely matched to the definition adopted in the VRAG calibration sample; a reconviction of attempted or completed homicide, assault, all sex crimes, armed robbery and forcible confinement. During the follow-up period, 29% of the sample was reconvicted of a violent crime.

The following H-scale items were significantly correlated with violent recidivism: H7 ($r = .42$; $p < .01$), H1 ($r = .36$; $p < .01$), H8 ($r = .20$; $p < .05$), H9 ($r = .28$; $p < .05$), H2 ($r = .24$; $p < .05$), H10 ($r = .38$; $p < .01$), and H5 ($r = .3$; $p < .01$). The AUC for the whole H-scale to predict violent recidivism was .76, while the AUC for the VRAG was .68.

The predictive validity of both the VRAG and the H-scale was considered to be moderate. There was a reported trend for the H-scale to perform slightly better compared to the VRAG. For the H-scale, most of the items had a positive correlation to recidivism and contributed well to the overall performance of the sub-scale. When similar items from the VRAG and the H-scale were compared several differences emerged.

Project and Scholarly Work

Urheim, R., Jakobsen, D., & Rasmussen, K. (2003, August). *Dimensions of inpatient aggressive behavior in a security ward: What is being "predicted"?* Paper presented at the 5th Nordic Symposium on Forensic Psychiatry, Ystad, Sweden.

Summary

This study examined the utility of the HCR-20 and PCL-R in predicting institutional violence in a secure psychiatric facility in Norway. Participants were 44 men (mean age = 31.8 years) and 7 women (mean age = 25.7 years) admitted over a 10 year period. Most participants had a primary diagnosis of schizophrenia (41%) or other psychosis (47%). The remaining 12% had a primary diagnosis of a personality disorder.

Mean HCR-20 scores were: Total (23.5, $SD = 6.8$); H-scale (13.8, $SD = 4.3$); C-scale (5.9, $SD = 1.9$); R-scale (3.9, $SD = 2.0$). Mean PCL-R scores were: Total (19.4, $SD = 8.8$); Factor 1 (7.6, $SD = 3.5$); Factor 2 (9.1, $SD = 5.2$).

The Staff Observation Aggression Scale was used to code aggressive episodes. Correlations were computed between the frequency of aggression (i.e., total episodes divided by patient days), severity of episode, and occurrence of physical aggression. The HCR-20 total score, C-scale, and R-scale (risk in institution) were correlated significantly with frequency (r s between .36 and .40), and there was a trend toward significance for the H-scale ($r = .28$, $p = .06$). The HCR-20 total score, H-scale, and R-scale were correlated significantly with severity (r s between .38 and .44). The HCR-20 was not correlated significantly with physical aggression ($r = .29$, $p = .08$); nor was the PCL-R (PCL-R F2, $r = .28$, $p = .08$). The only significant correlations obtained with the PCL-R were between frequency and total score and Factor 2 and between severity and total score. Frequency

and proportion of physical aggression were significantly higher among women than men.

AUC values for the prediction of frequency were: HCR-20 total (.76); H-scale (.67); C-scale (.82); Risk (.70); PCL-R Factor 1 (.64); PCL-R Factor 2 (.77). AUC values for the prediction of most severe episode were: HCR-20 total (.82); H-scale (.77); C-scale (.73); Risk (.76); PCL-R total (.73); PCL-R Factor 1 (.65); PCL-R Factor 2 (.71).

Project and Scholarly Work

Vincent, G. M. (1998). *Criminal responsibility after Bill C-30: Factors predicting acquittal and lengths of confinement in British Columbia*. Unpublished master's thesis, Simon Fraser University, Burnaby, British Columbia, Canada.

Summary

This is a chart review study of 250 persons referred from court to a maximum security forensic institute over the course of five years for the purpose of assessment of criminal responsibility. The focus of this summary is the 125 persons who were found Not Criminally Responsible on Account of Mental Disorder (NCRMD) for their offences. The sample (M age = 34.98; $SD = 10.67$) was primarily male (82.4%), Caucasian (77.4%), single (88.6%), unemployed (76.4%), and many patients had less than grade 11 education (40.2%). Most patients had committed a violent index offence (77.6%), and most had a primary diagnosis of a psychotic disorder at assessment (66.9%), followed by mood disorder (21.0%)

The purpose of the study was to evaluate which factors predicted (1) verdicts of NCRMD (insanity acquittal) versus guilty, (2) length of confinement and days in the system. The HCR-20 was used for the latter, as a predictor of days in the system and days confined. Also included in such analyses were a variety of criminological, psychiatric, demographic variables, and the Psychopathy Checklist: Screening Version. Hierarchical Cox Proportional Hazards Regression was used as the method of prediction, with time in the system as the dependent measure. After all blocks and variables were entered, the HCR-20 was the only significant predictor, with an $e^B = .898$ (Odds = 2.45). In particular, the R Scale was the strongest of the three scales. Using a somewhat more liberal approach with a backward elimination entry procedure, one other variable in addition to the HCR-20 entered the equation (offence severity). For a slightly different dependent measure (days until first release), several variables entered the model (using backward elimination): level of violence, number of remand charges, homicidal at offence, age at first mental health contact, PCL:SV, and HCR-20.

Months confined, in the system, and until first release were calculated as a function of low, moderate, and high scores on the HCR-20 (by dividing the total scores into thirds). Months in the system, confined, and until first release, for the LOW group were 32.82, 9.22, and 7.25, respectively. For the MODERATE group, results were 38.68, 18.56, and 13.93, respectively. For the HIGH group, results were 45.47, 40.23, and 30.92, respectively. These findings provide support for the concurrent validity of the HCR-20. The factors it predicted are related to legal concepts of risk and threat.

Project and Scholarly Work

Vincent, G. M., Ross, D. J., Whittemore, K., Eaves, D., Hart, S. D., Ogloff, J. R. P., & Webster, C. D. (2001, April). *Using the HCR-20: File-based researcher ratings vs. file + interview-based clinician ratings*. Paper presented at the Founding conference of the International Association of Forensic Mental Health Services, Vancouver, BC, Canada.

Summary

Vincent et al. (2001) investigated the correspondence between HCR-20 (Version 1) C and R scale ratings made by psychiatrists from file + interview, and version 2 ratings made by researchers from file alone. There was a significant difference between clinicians and researchers on the C scale, but not on the R scale. The difference, though significant, was small (Cohen's $d = .31$). The association between rater groups was fairly high ($ICC_1 = .58$ for C scale; $ICC_1 = .70$ for R scale). There were few (0-4 per item) "0-2" coding disagreements on individual items. Vincent et al. (2001) concluded that the file-based ratings were sufficiently reliable for research purposes.

Project and Scholarly Work

Watt, A., Topping-Morris, B., Rogers, P., Doyle, M., & Mason, T. (2003). Pre-admission nursing assessment in a Welsh medium secure unit (1991-2000): Part 2 – comparison of traditional nursing assessment with the HCR-20 risk assessment tool. *International Journal of Nursing Studies*, 40, 657-662.

Summary

The authors noted that pre-admission forensic nursing assessment does not have an empirical evidence base despite the fact that such assessments are completed routinely and at substantial financial cost. The purpose of this study was to assess retrospectively the quality of pre-admission risk assessments completed by nurses at one forensic psychiat-

ric clinic in Wales through comparison with the HCR-20, Version 1.

Eighty-five consecutive cases referred for forensic nurse assessment over a 51-month period were coded as to whether sufficient information had been collected during the risk assessment to allow scoring of the HCR-20 items. HCR-20 item 7 (PCL-R score) was omitted because the researchers did not have formal training in administration of the PCL-R. In roughly 89% of cases, there was sufficient information to rate the 5 items comprising the C-scale. Pertaining to the H-scale, 7 of the 9 items could be rated in over 80% of the cases. Presence of a personality disorder and employment history could be rated in 71% and 77% of the cases, respectively. Four of the five R-scale items could be coded for over 85% of the cases. The Risk management item that assesses the feasibility of plans was able to be rated for less than 80% of the cases.

Mean scores were: HCR-20 Total 23.6 (3.6); H-scale 12.5 (2.8); C-scale 5.5 (3.4); R-scale 5.5 (3.1). In the discussion section the authors concluded that information traditionally collected by forensic nurses in the course of a risk assessment was appropriate, but noted the importance of research-based practice.

Project and Scholarly Work

Whittemore, K. E. (1999). *Releasing the mentally disordered offender: Disposition decisions for individuals found unfit to stand trial and not criminally responsible*. Unpublished Ph.D. Dissertation, Simon Fraser University, Burnaby, British Columbia, Canada.

Summary

This is a chart review study of 172 persons either found unfit to stand trial ($n = 50$) or not criminally responsible on account of mental disorder (NCRMD; $n = 122$). The sample (M age = 37) was primarily male (83.14%), Caucasian (80.23%), single (88.37%), and unemployed (79.07%). Most patients had committed a violent index offence (75.58%), and most had a primary diagnosis of a psychotic disorder at assessment (60.47%), followed by bipolar disorder (16.28%).

The purpose of the study was to evaluate which factors predicted criminal review board release decisions (discharge versus custodial detention). A variety of mental health, criminological, and demographic characteristics were used as predictor measures along with the HCR-20 and PCL:SV. Hierarchical logistic regression was used as the method of prediction, with release decision as the dependent measure. For the first review board hearing (patients have regular hearings until released), the H Scale, C Scale, and R Scale were entered in separate blocks. Each

END OF FORENSIC PSYCHIATRIC STUDIES

was a significant predictor (lower scores relating to discharge), with C and R adding incrementally to H. R was most strongly related to discharge decisions. Additional analyses were carried out to predict subsequent discharge/custody decisions. Three Clinical subscale items (Negative Attitudes, Lack of Insight, and Impulsivity), and one Risk Management scale item (Noncompliance with Remediation Attempts) predicted discharge.

These results suggest that at the first hearing, the Risk Management items were most important for discharge decisions, although the Clinical and Historical items also were predictive. At subsequent hearings, change in mental status (Clinical Scale items) emerged as the more important predictor. Results provide support for the concurrent validity of the HCR-20. Release decisions legally require the Review Boards to take into account the threat posed by the individual, the need to reintegrate the accused into society, and the mental condition of the accused.

Project and Scholarly Work

Wintrup, A. (1996). *Assessing risk of violence in mentally disordered offenders with the HCR-20*. Unpublished master's thesis, Simon Fraser University, Burnaby, British Columbia, Canada.

Summary

This is a chart review study of 80 men remanded to a secure forensic facility. The mean age at admission was 32.6 years ($SD = 10.8$). The majority of patients had been previously hospitalized in a psychiatric setting (77.5%), and most had previous charges or convictions for criminal offences (78.9%). Both the HCR-20 and the PCL-R averaged correlations just below $r = .30$ with several measures of later community violence. The HCR-20 was quite strongly related to subsequent re-admissions to the forensic hospital ($r = .38$) and to psychiatric hospitalizations ($r = .45$). The relationship of the PCL-R to these same outcomes was not as strong, at $r = .25$ and $r = .36$, respectively. However, whether these re-hospitalizations involved violence was not specified.

See also

Douglas, K. S., Webster, C. D., & Wintrup, A. (1996, August). The HCR-20 risk assessment scheme: Psychometric properties in two samples. Poster presented at the annual convention of the American Psychological Association, Toronto.

CORRECTIONAL SETTINGS

(INCLUDES MENTALLY DISORDERED OFFENDERS)

Project and Scholarly Work

Belfrage, H., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: A prospective study in two maximum-security correctional institutions. *Journal of Forensic Psychiatry, 11*, 167-175.

Summary

The HCR-20 violence risk assessment scheme (Version 2) and the Psychopathy Checklist: Screening Version (PCL:SV) were coded on a sample of 41 male inmates from two Swedish maximum security prisons. The two coders were Ph.D. and M.D. level clinician-researchers. This was a prospective study of violence within the correctional institution over an eight month period. The HCR-20 and PCL:SV were coded by use of both file review and clinical interview. The R Scale of the HCR-20 was coded using the "In" option as explained in the manual (Version 2).

The mean age of the participants was 35, and the mean length of incarceration at time of assessment was three years. All participants had a personality disorder (mostly antisocial). Of the 41, 27 were incarcerated for homicide, and 14 for other violent offences. The sample was highly psychopathic, with 30 of 41 inmates being classified as psychopaths.

Eight of the 41 (19.5%) inmates were violent in the prison. The C Scale, R Scale, HCR-20 Total Score, PCL:SV Part 2, and PCL:SV Total Score differentiated between the violent and non-violent groups. The HCR-20 Total score was 33.4 in the violent group, and 24.6 in the non-violent group. All HCR-20 R Scale items were significantly greater among the violent group than the non-violent group. The H Scale was not predictive of violence, except for Item H10. In the group of 30 psychopaths, the R Scale and HCR-20 Total score were significantly higher in the violent inmates. Four of the five R Scale items were higher in the violent psychopaths compared to the non-violent psychopaths.

Results imply that the HCR-20 (and the PCL:SV) are predictive of violence by inmates within correctional institutions. Even among a sub-group of psychopaths, the HCR-20 distinguished between violent and non-violent inmates. The authors comment that the H Scale was not predictive in this sample because inmates (being maximum security

violent inmates) were homogeneous with respect to most historical factors. The Clinical and Risk Management factors did, however, provide for a means of separating violent from non-violent inmates. These results are consistent with those of Strand et al. reported above. The results of the study, though limited by a small sample, provide support for the importance of risk management concerns for high-risk violent offenders.

Project and Scholarly Work

Belfrage, H., Fransson, G., & Strand, S. (2004). Management of violent behavior in the correctional system using qualified risk assessments. *Legal and Criminological Psychology, 9*, 11-22.

Summary

This research focused on whether the use of comprehensive risk assessments and subsequent risk management could prevent institutional violence in a maximum security correctional institution in Sweden. More specifically, the authors investigated whether increasing staff members' knowledge of risk factors and appropriate risk management strategies would affect the rate of violence on a ward.

All participants ($N = 47$) were offenders with a violent criminal history and who were incarcerated at some time between October 1999 and June 2002 on one ward of the institution. The authors noted that this particular ward (A-ward) is not meant to house any specific type or class of offender and that it is not known to have an elevated level of psychiatric problems among prisoners relative to other wards. However, many of individuals housed on A-ward were transferred there for misbehaviour. Participants on average were 32 years old and most frequently were diagnosed with antisocial personality disorder ($n = 29$). No participants were diagnosed with a thought disorder. Eighteen participants had PCL:SV scores of 18 or above, and psychopathic offenders had significantly higher HCR-20 scores compared with nonpsychopathic offenders ($p = .000$, Mann-Whitney U -test).

Prior to the study's commencement, staff received training on risk assessment, common risk factors for violence, foundational knowledge about the HCR-20 and PCL:SV, and adequate risk management strategies for different types of mentally disordered offenders. Ongoing training also was provided throughout the study. Two clinicians collaboratively completed a risk assessment on each participant that included the Swedish versions of the HCR-20 and the PCL:SV. Following completion of the risk assessment, results were discussed with staff members and a risk management strategy for the individual was developed.

The HCR-20 was readministered to roughly one third ($n = 13$) of the sample. The average length of time between the two HCR-20 assessments was 12 months (range: 3-24 months). Comparison of HCR-20 scores between the follow-up group of 13 and the entire study group of 47 revealed no significant changes. However, there was a significant decrease in incidence of violence (from an average of 14 violent incidents per year between 1993-1998 to an average of 5 violent incidents per year during the study period of 1999-2001; $p = .024$, Mann-Whitney U -test). This 64% reduction is in stark contrast to all other wards in the prison, in which there were no decreases in incidence of violence during the study period.

The discussion section advances possible explanations for the results. The authors highlight the importance of incorporating into a risk assessment protective factors, which they noted can reduce violence even when important risk factors do not decrease.

Project and Scholarly Work

Cooke, D. J., Michie, C., & Ryan, J. (2001). *Evaluating risk for violence: A preliminary study of the HCR-20, PCL-R and VRAG in a Scottish Prison Sample*. Report prepared for the Scottish Prison Service.

Summary

This was a retrospective follow-up of 250 male adult prisoners released from a Scottish prison who had been randomly selected from the larger population. Mean age was 26.8; almost all were Caucasian; roughly half were unmarried (46%); offences included drugs (36%), assault (24%), theft and break and enter (20%), homicide (8%), weapons (7%), kidnapping (3%), and sexual offences (2%). Prisoners had received earlier comprehensive correctional assessments, using interview and file procedures. Psychopathy was one of the constructs that was assessed. The HCR-20 later was coded from files. The researchers were unable to code C1, C5, R2, R4, and R5. As they commented, this may have underestimated the predictive accuracy of the HCR-20. Violence was measured with the MacArthur Community Violence Instrument. The authors carried out comprehensive analyses of the HCR-20, VRAG, and PCL-R, down to the item (this is a 128 page report!).

Interrater reliability, using ICC_1 on a subsample of 60 prisoners, was described as “acceptable and ... consistent with values obtained in other studies” (p. 30). Values were as follows: HCR-20 Total Scale Score (.92); Historical Scale Score (.92); Clinical Scale Score (.74); Risk Management Scale Score (.70). Lower indices for C and R may stem from coding files only.

Survival analyses of individual HCR-20 items led the researchers to conclude that “Overall, these analyses indicate that the vast majority of these items have some potential utility as predictors” (p. 46). Survival analyses indicated that the Total H Scale Scores were related to each outcome measure (including both violent and non-violent convictions); the C Scale was related to the two violence outcome measures, and the R Scale was related to the general recidivism outcome measures, but only weakly to violent outcomes. Recall that only two of five R Scale items were coded.

Cox proportional hazard model analyses revealed that the e^B for reincarceration for violence was 1.20 (exponentiated value of the model parameter; an effect size indicating the change in hazard rate as a function of a unit change in the total HCR-20 score). This means that for every 1-point increase in the HCR-20, the hazard for violence increases 20%. For a 5-point increase, the hazard increases by 200% ($100\% + 5 \times 20\%$). Using the H Scale alone, these analyses revealed that 14% of those with a mean score (10.9) were returned to prison for violence, 4% with a score of 5, and 72% with a score of 20. This was greater discrimination than achieved for the PCL-R or VRAG. When compared directly, the H Scale was the strongest predictor of return to prison for violence; the VRAG was stronger for general recidivism and violent recidivism that did not lead to reimprisonment (presumably less serious violence). Cooke et al., based on these and other analyses, interpreted the findings as suggestive of greater specificity for the prediction of violence, and likely serious violence, compared to general recidivism, for the H Scale compared to the VRAG and PCL-R. AUC values for community outcomes were not significantly different for the various measures and outcome criteria (averaging approximately .70). HCR-20 AUC s ranged from .69 to .74; VRAG AUC s from .67 to .73; PCL-R AUC s from .65 to .72.

Institutional violence also was studied, with some different findings emerging. Many of the HCR-20 items predicted violence. The e^B values for the HCR-20 were as follows: HCR-20 Total Scale Score (1.12); H Scale Score (1.16); C Scale Score (1.45); R Scale Score (1.09). For the PCL-R, e^B was 1.06, and for the VRAG, 1.08. As such, for a 1-point increase on the H-Scale, the hazard for violence increases 16%; for a 1-point increase on the C Scale, the hazard increases 45%. Multivariate Cox proportional hazard model analyses showed that there not meaningful differences between the various measures, although, as with community violence, the H Scale had somewhat greater specificity than the VRAG and PCL-R in its relationship to violent institutional infractions, rather than all infractions per se. AUC values were moderate for all predictors across general and violent institutional infractions, and did not differ from one another (HCR-20 = .64 - .64; H Scale = .64 - .65; PCL-R = .61 - .63; VRAG = .66 - .67).

Although there were not substantial differences between the measures in terms of predictive validity, with the HCR-20 perhaps demonstrating greater specificity than the PCL-R and VRAG vis-à-vis serious versus minor offending and violence, Cooke et al. concluded that the HCR-20 “remains the instrument of choice because it provides guidance on how to *manage* risk not merely how to *predict* risk” (p. 3, Executive Summary).

Project and Scholarly Work

Dahle, K-P. (2001). Unpublished data analyses. Personal Communication. January 17, 2001.

Summary

This study used the German version of the HCR-20. It drew from data from the Berlin CRIME study, which was a longitudinal study of 397 criminal offenders released from prison in 1976. The HCR-20 was coded on 200 randomly selected prisoners from these 1976 files. The researchers were unable to code H7 – the Psychopathy item. Because the sample was random, index offences mainly were property offences, though there were some serious violent offences as well. The inmates averaged 31.42 ($SD = 5.40$) years in age, and had 6.46 ($SD = 4.72$) previous convictions. More than half (65%) had committed previous violent offences. Many had alcohol (56%) problems; fewer (17%) had drug problems.

Based on a subsample of 30 offenders, interrater reliability (Kendall’s Tau) was .80 for the HCR-20 Total Scale. Spearman’s Rho was .731 for the C Scale and .930 for the H Scale. Interrater reliability was not reported for the R Scale. Over the 20 year follow-up, the correlation between the HCR-20 and violent recidivism was as follows: HCR-20 Total Scale Score (.25); H Scale Score (.24); C Scale Score (.23); R Scale Score (.10).

Project and Scholarly Work

Douglas, K. S., Yeomans, M. & Boer, D. (in press). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior*.

Summary

This study compared the predictive validity of five indices of violence risk – the HCR-20, the VRAG, the VORAS, the PCL:SV, and the PCL-R. Participants were 188 male offenders released from federal corrections institutions to supervision in Western Canada. Participants were selected based on known outcome status after release. This status was as follows: violent recidivism ($N = 93$) or no violent

recidivism after release ($N = 95$). The follow up period for this study ranged from 6 to 11 years.

Inter-rater reliability was good to excellent for all measures except for the HCR-20 structured final judgment ratings which can be considered fair to moderate. Even with the lower inter-rater reliability, there were no low/high disagreements with this item.

Point-biserial correlations between risk assessment measures and violent recidivism showed that with the exception of the H-scale, the HCR-20 total and sub-scales produced correlations of approximately .50. The aforementioned H-scale showed a correlation of .36. The AUCs for the HCR-20 were approximately .80 (up to .82) except for the H-scale (.72). Partial point-biserial correlations were also conducted with the HCR-20 total score, with the correlation to violent recidivism dropping from .51 to .25 after controlling for the VRAG, PCL-R and the VORAS. Other measures also showing a significant positive point-biserial correlation after controlling for other measures were: HCR-20 C and R scales and structured final judgment, the VRAG, VORAS B, and Cooke and Michie’s third factor. The other measures were either not significantly related or were negatively related to the outcome.

Binary logistic regressions were conducted to directly compare the measures. Using forward conditional entry procedures only the HCR-20 total score entered the model. Using direct entry procedures the HCR-20 total score, VRAG total score, and the VORAS total score entered the model. Using subscales instead of total scores, direct entry showed that the HCR-20 C-scale, VORAS A (negatively) and VORAS B were significant predictors. Using forward entry, the HCR-20 C-scale, VORAS A (negatively) and VORAS B were again significant predictors. Using the HCR-20’s and other measures’ final risk judgments showed that the HCR-20 structured clinical final risk judgment, the VRAG’s actuarial categorical system, and the VORAS actuarial final risk score were significantly predictive.

The discussion section explores the implications of the results from this study in regards to the use of the five measures analyzed here.

Project and Scholarly Work

Douglas, K. S. & Webster (1999). The HCR-20 violence risk assessment scheme: Concurrent validity in a sample of incarcerated offenders. *Criminal Justice and Behavior*, 26, 3-19.

Summary

The HCR-20 violence risk assessment scheme was coded in a sample of 72 Canadian, male, federally-sentenced,

maximum security offenders who had been referred to a regional health centre of the Correctional Services of Canada. The concurrent validity of the HCR-20 was assessed through comparison to other instruments (the Psychopathy Checklist - Revised; Violence Risk Appraisal Guide) and to the presence of several past indexes of violent and antisocial behavior. Only the H and C scales could be coded because no offenders had yet been released. This was a postdictive study.

The interrater reliability of the H and C combined scores was .80. Correlations between the number of previous violent charges and the H scale, C scale, and their combination ranged from moderate to large.¹ The Historical scale correlated at $r = .50$ with previous violence (with the "previous violence" item removed from the H scale), the Clinical scale at $r = .30$, and the combined total at $r = .44$. The VRAG correlated at $r = .20$ with previous violence, and the PCL-R's correlation with past violence was $r = .41$.

Scores above the median of the HCR-20 increased the odds of the presence of various measures of past violence and antisocial behavior by an average of four times. The main limitations of this research were a small sample and a retrospective design.

See also

Douglas, K. S., Webster, C. D., & Wintrup, A. (1996, August). *The HCR-20 risk assessment scheme: Psychometric properties in two samples*. Poster presented at the annual convention of the American Psychological Association, Toronto.

Project and Scholarly Work

Doyle, M., Dolan, M., & Mc Govern, J. (2002). *The validity of North American risk assessment tools in predicting in-patient violent behaviour in England*. *Legal and Criminological Psychology*, 7, 141-154.

Summary

This study compared the validity of the Historical component of the HCR-20, the Psychopathy Checklist: Screening Version (PCL:SV; Hart, Cox, & Hare, 1995) and the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993). This research used 87 adult mentally disordered patients in a medium secure unit. An independent researcher, who was blind to the score on the risk assessments, collected data on violent outcome measures during the first 12 weeks after admission for all 87 subjects.

For this study, violence was operationally defined as actual, attempted or threatened harm to others. Violence episodes were dichotomized into two levels. Level 1 involved physical assault against a person or any violence resulting in injury to a person. Level 2 included other aggressive acts such as threats or property damage.

On the risk assessment tools, those with any violent outcome had significantly higher scores on all measures, with the PCL:SV showing the most significant group difference ($t = 4.77$; $p < .0001$). The VRAG ($t = 3.6$; $p < .001$) and the H-10 ($t = 3.2$; $p < .001$) also showed a significant group difference.

Using ROC analyses for the prediction of any type of violence, the AUC's produced for the PCL:SV total score ranged from .76 ($p < .001$) (for any and physical violence) to .74 ($p < .01$) (for level 1 violence). The AUCs produced for the VRAG total score ranged from .71 ($p < .01$) (for any and physical violence) to .64 ($p < .01$) (for level 1 violence) and the AUCs produced for the H-10 total score ranged from .70 ($p < .01$) (for any and physical violence) to .66 ($p < .01$) (for level 1 violence).

Using a cut-off score of 18 (recommended manual cut-off) on the PCL:SV, the odds ratio for any violence was 4. Using a cut-off score of 5 (sample mean) on the VRAG, the odds ratio was 3.75 for any violence. Using a cut-off score of 12 (sample mean) on the H-10, the odds ratio was 3 for any violence.

The PCL:SV total score correlated significantly with the VRAG score ($r = .81$, $p < .001$) and with the H-10 score ($r = .8$, $p < .01$). The VRAG score and the H-10 score also correlated significantly with each other ($r = .83$, $p < .001$).

Forward stepwise logistic regression showed that only the PCL:SV total score contributed significantly to the prediction of any violence ($\chi^2 = 20.05$, $p < .001$).

Implications for research on risk assessment, as well as the clinical assessment and management of violence, are discussed.

Project and Scholarly Work

Dunbar, E. (2003). *Symbolic, relational, and ideographical signifiers of bias-motivated offenders: Toward a strategy of assessment*. *American Journal of Orthopsychiatry*, 73, 203-211.

Summary

This study was not a validation study of the HCR-20, but rather an analysis of the characteristics of hate-motivated violent offenders in California, and comparison of this group to other offenders. The HCR-20 was used as a meas-

¹ According to Cohen (1992), a moderate size correlation is $r = .30$, and a large correlation is $r = .50$.

ure of violence risk for this purpose. The HCR-20 was coded from files of 58 hate-motivated criminals. Most of the offenders were male ($n = 53$, 91.4%), with a mean age of 24.5 ($SD = 8.07$). Most offenders were Latino (48.3%), followed by Euro-Caucasian (32.8%) and African American (15.5%). Close to half were unemployed (45.5%). The majority of the offenders had substance use problems (58.6%), and many (22.4%) had received past psychiatric treatment. The vast majority of offenders had previous convictions (87%) and a history of violence (60%). All offenders had violent, hate-related index offences.

Analyses on the HCR-20 included correlations with the PCL-R, with violent offences, and with the Cormier-Lang crime severity scales. The HCR-20 did not correlate with the seriousness of the index offence (although there is no clear reason to expect it to). Correlations between the HCR-20 total score and indices of past crime and violence ranged from .33 to .63. For the H scale, the range was .39 to .68, for the C scale, .30 to .56, and for the R scale, .19 to .45. In general, the correlations were above .40, and many were above .50.

This study is consistent with findings from Canadian violent offenders (see Douglas & Webster, 1999, above) in terms of the relationship with violence. Although the study included post-dictive analyses (as did Douglas & Webster, 1999), the effects are generally large, often exceeding .50, and ranging to .68. The findings support the concurrent validity of the HCR-20, and support the effort of doing larger scale research on the HCR-20 in American criminal offenders.

Project and Scholarly Work

Gray, N. S., Hill, C., McGleish, A., Timmons, D., MacCulloch, M. J., & Snowden, R. J. (2003). Prediction of violence and self-harm in mentally disordered offenders: A prospective study of the efficacy of the HCR-20, PCL-R, and psychiatric symptomatology. *Journal of Consulting & Clinical Psychology, 71*, 443-451.

Summary

The authors investigated prospectively the predictive validity of the HCR-20 (Total, H-scale, and C-scale), PCL-R, Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974), and Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962) with respect to institutional misconduct and self-harm/suicide. The R-scale was not scored because evaluations were completed within 2 weeks of admission at which time participants' social situations at discharge were difficult to project. The sample consisted of 34 mentally disordered offenders admitted to one of two medium-secure hospital units in the U.K. All participants had committed a serious offense (e.g., manslaughter, sexual of-

fense, physical assault, arson). Almost half were diagnosed with paranoid schizophrenia (44.1%); 23.5% depression; 14.7% personality disorder; 17.6% 'other,' such as bipolar disorder, organic disorder, neurosis). At admission, the mean age was 33 years ($SD = 11.9$). Participants primarily were male (76.5%) and White (85.3%).

Incidents relating to four outcome criteria were measured: (1) verbal aggression, (2) aggression to property, (3) physical aggression to staff or other patients, and (4) self-harm or suicide. All incidents were assessed using a structured record form developed for this study (the Aggression Vulnerability Scale; AVS). The authors reported good intraclass correlations for the AVS for all four outcome criteria in a pilot study. AVS ratings were completed weekly for 3 months (or until the participant was transferred off the unit) using nursing records, incident report forms, and weekly interviews with the primary nurses.

The mean scores for the measures were as follows. H-scale (10.2; $SD = 3.4$; range 3-18); C-scale (5.4; $SD = 2.6$; range 1-10); PCL-R total (10.4; $SD = 1.5$; range 0-33); Factor 1 (3.9; $SD = 4.3$); Factor 2 (5.5; $SD = 4.2$); BHS (8.6; $SD = 6.4$; range 1-19); BPRS (42.0; $SD = 2.7$; range 17-74). Correlations between the measures tended to be high and significant. Notably, there was an especially strong relationship between the BPRS and the HCR-20 total score (.63) and C-scale (.71).

The authors used correlations, ROC analysis, odds ratios, and the Mann-Whitney U Test (a nonparametric comparison of means for those scoring higher versus those scoring lower than the median value) to examine predictive validity for the HCR-20, PCL-R, BHS, and BPRS with respect to each of the four dependent measures. For verbal aggression and HC total score: $r = .53$, $p < .001$; AUC = .79, $p < .001$; OR = 2.55; Mdn split $p = < .01$. For verbal aggression and H-scale: $r = .43$, $p < .01$; AUC = .73, $p < .05$; OR = 2.21; Mdn split $p = < .01$. For verbal aggression and C-scale: $r = .49$, $p < .01$; AUC = .74, $p < .01$; OR = 2.27; Mdn split $p = < .05$. For violence to property and HC total score: $r = .56$, $p < .001$; AUC = .83, $p < .001$; OR = 8.85; Mdn split $p = < .01$. For violence to property and H-scale: $r = .54$, $p < .001$; AUC = .82, $p < .001$; OR = 8.45; Mdn split $p = < .01$. For violence to property and C-scale: $r = .49$, $p < .001$; AUC = .77, $p < .001$; OR = 3.85; Mdn split $p = .09$. For physical aggression and HC total score: $r = .53$, $p < .001$; AUC = .81, $p < .001$; OR = 8.25; Mdn split $p = < .01$. For physical aggression and H-scale: $r = .43$, $p < .01$; AUC = .77, $p < .001$; OR = 7.46; Mdn split $p = < .05$. For physical aggression and C-scale: $r = .49$, $p < .01$; AUC = .79, $p < .001$; OR = 7.42; Mdn split $p = < .05$. The HCR-20 was not predictive of self-harm, which is not surprising given that it was developed to predict risk of violence to others.

Values for the predictive indices yielded for the HCR-20 were larger than those for the PCL-R. The PCL-R total score was, however, a useful predictor for violence to property and physical violence ($r_s = .35 - .38$; AUCs = .70-.76; ORs = 1.88 – 2.84). The PCL-R was not a significant predictor of verbal aggression or self-harm.

Project and Scholarly Work

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, 28, 471-489.

Summary

Kroner and Mills (2001) completed a prospective study of institutional misconduct among offenders. At intake, they completed the HCR-20, along with other instruments, on a sample of 97 consecutively admitted Canadian federal offenders. In terms of interrater reliability, ICC_1 was reported to be .85 for the total score. The HCR-20 correlated .32 with minor misconducts, and .11 with major misconducts, and was not significantly different than the PCL-R, the VRAG, or the LSI-R. For post-release analyses, the HCR-20 correlated at .28, .16, .21, and .39 with total convictions, violent convictions, nonviolent convictions, and revocations, respectively. Again, there were no significant differences between measures.

It should be pointed out that, although there were no significant differences between measures, that the coding procedure was not optimal for community violence. The instruments were coded at admission, and this score was used to predict violence after release from the institution, some years later. This has particularly strong implications for measures that will change over time, such as the C and R scale and the LSI-R. Further, institutional outcome criteria (misconduct) included mostly *non-violent* indices, such as improper dress, disrespect, noncompliance with directions, drug use, and refusing urinalysis.

Kroner and Mills (2002/personal communication). Unpublished supplement for Kroner and Mills (2001), above.

Summary

Authors reported correlations that were not published in their original report, describing the relationship between risk measures (HCR-20 and VRAG) and violent re-offending over a longer period than published in the article.

Significant correlations between the HCR-20 items and violent behavior are as follows. Total score ($r = .37$; $p <$

.01), H2 ($r = .28$; $p < .05$), H3 ($r = .28$; $p < .05$), H4 ($r = .23$; $p < .05$), H8 ($r = .22$; $p < .05$), C1 ($r = .27$; $p < .05$), C2 ($r = .34$; $p < .01$), C4 ($r = .28$; $p < .05$), C5 ($r = .25$; $p < .05$), R1 ($r = .26$; $p < .05$), R2 ($r = .26$; $p < .05$), R4 ($r = .35$; $p < .01$) and R5 ($r = .27$; $p < .05$).

Significant correlations between the VRAG and violent behavior are as follows: Total score ($r = .28$; $p < .05$), item #2 ($r = .29$; $p < .01$), item #6 ($r = .28$; $p < .05$) and item #11 ($r = .27$; $p < .05$).

Project and Scholarly Work

Logan, C., Doyle, M. & Dolan, M. (2001, November). *Risk assessment in English male offenders who have committed single acts of serious violence or sexual violence: how can the past be used to predict the future?* Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together, Sundsvall, Sweden.

Summary

This research studied the differences on the HCR-20, VRAG and PCL-R between groups of men who committed single acts of violence vs. men who have a history of committing multiple acts of violence. The authors point out that there may be difficulty in predicting future actions of men who have only committed one single act of violence due to the fact that that risk assessment research is predicated on the assumption that static indicators of past behavior can be used to predict future behavior. Subjects were male mentally disordered offenders, 21 years of age or older with one or more convictions for violence or sexual violence.

In terms of the HCR-20, there were differences between groups on many individual items from the three subscales. On the H-scale, there were significant differences on H2, H7, H8, H9, and H10, with repeat offenders scoring higher on all of these items. On the C-scale, there were significant differences on C2, C3 and C5, with repeat offenders scoring higher on C5 and C2 and lower on C3. On the R-scale, there were significant differences on R1, R2, R3 and R4, with repeat offenders scoring higher on all of these items. In terms of the PCL-R, repeat offenders scored significantly higher on the total score and on the factor two scale. In terms of the VRAG, repeat offenders scored significantly higher on the total score.

Project and Scholarly Work

Logan, C. & Watt, K. (2001, November). *Structured professional guidelines approaches to risk assessment: Single practitioner vs. multidisciplinary team administration*. Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden.

Summary

This research project investigated the comparability of outcomes when the HCR-20 was given in single administrations or in group administrations. This study sampled from a group of 85 male mentally disordered offenders in a personality disorder unit of a major forensic psychiatric facility in England. Descriptive statistics for group administrations were as follows: HCR-20 Total score, $N = 68$, $M = 27.88$ ($SD = 5.01$); H-score, $N = 73$, $M = 15.15$ ($SD = 2.53$); C-scale, $N = 84$, $M = 5.27$ ($SD = 2.2$) and R-scale, $N = 80$, $M = 7.5$ ($SD = 2.27$). Descriptive statistics for single administrations were as follows: HCR-20 Total score, $N = 36$, $M = 27.33$ ($SD = 4.76$); H-score, $N = 36$, $M = 14.75$ ($SD = 2.45$); C-scale, $N = 36$, $M = 5.53$ ($SD = 1.95$) and R-scale, $N = 36$, $M = 7.06$ ($SD = 1.67$). None of these scores for the group or single administrations were significantly different from each other.

Correlations (ρ) between group and single administrations were mostly significant. HCR-20 total ($r = .428$, $p < .05$), H-scale ($r = .654$, $p < .01$) C-scale ($r = .503$, $p < .05$) and R-scale (ns).

The authors concluded that if there are no differences between single and group administrations of the HCR-20, then why not make use of group administrations.

Project and Scholarly Work

Pham, T. H., Claix, A., & Remy, S. (2000, June). *Assessment of the HCR-20 in a Belgian prison sample*. Paper presented at the 4th European Congress on Personality Disorders. Paris, France.

Summary

This study collected HCR-20 and PCL-R data from 68 offenders in a Belgian high security prison. The mean follow-up period, available for 38 offenders, was 1010 ($SD = 894$) days. Interrater reliability, based on a subsample of 10 offenders, was .85 (Pearson r) for the HCR-20 Total Scale Score. The AUCs between HCR-20 indices and violent recidivism were as follows: HCR-20 Total Scale (.76); Historical Scale (.77); Clinical Scale (.74); Risk Manage-

ment (.71). PCL-R AUCs were .82 (Full Scale), .77 (Factor 1), and .75 (Factor 2). Based on a subset of 20 offenders, correlations between the HCR-20 indices and the Buss and Perry Aggression Questionnaire ranged from .29 to .57, and with the Heilbrun Index of Dangerousness in the Community, from .32 to .37.

END OF CORRECTIONAL STUDIES

MIXED SETTINGS

Project and Scholarly Work

Côté, G. (2001, April). *Violent behaviour, PCL-R and HCR-20 among involuntary inpatients, forensic patients and severely mentally disordered inmates*. Paper presented at the First Annual Meeting of the International Association of Forensic Mental Health Services, Vancouver, BC.

Summary

Côté observed that much research on the HCR-20 has been carried out in very different settings (i.e., civil psychiatric; forensic psychiatric; severe offenders; general offenders) and set out to evaluate the performance of the HCR-20 as a function of such settings. They sampled from involuntarily committed patients, forensic patients who had been found not criminally responsible on account of mental disorder, and mentally disordered inmates, all of whom resided in institutions throughout the Canadian province of Quebec. Although the study was prospective, the data reported in this presentation were postdictive. Côté used the French translation of the HCR-20. Participants ($n = 79$ after attrition and refusal factored in, 68 with criminal record data) were evaluated just prior to release into the community. Although both males and females were sampled, the small number of females prompted Cote to drop them from analyses for this presentation.

Interrater reliability, based on a subset of 20 cases, was reported for the H and C scales, respectively, as follows: $ICC_1 = .88$, $ICC_2 = .93$; and $ICC_1 = .71$, $ICC_2 = .83$. Alpha was reported to be .93 and .83, respectively, for H and C.

The H scale (with H1 removed) varied significantly across groups with no previous offences ($M = 8.43$; $SD = 3.10$), non-violent offences ($M = 9.17$; $SD = 3.64$) and violent offences ($M = 13.06$; $SD = 3.36$). The Cohen's d between the violent group and non-violent offence group was large ($d = 1.11$), as it was for the difference between the violent group and the no offence group ($d = 1.43$). The difference between C scores across these groups was not significant, although the effect was large for the violent versus no offences comparison ($d = .73$) (it was small – .29 – for the violent offences versus non-violent offences comparison).

The H scale, without H1, also was larger among correctional inmates ($M = 14.32$, $SD = 2.87$) than forensic patients ($M = 11.42$, $SD = 3.50$), or involuntary patients ($M = 10.32$, $SD = 3.76$) – which is consistent with other research when compared across studies. Côté reported multiple

comparison correlation coefficients (eta) of .54 for the H scale, .24 for the C scale, and .50 for the H scale without H1 with respect to offence group, and .4 (H scale) and .16 (C scale) within the legal status groups.

Côté reported AUC values, using Statistics Canada's definition of violence, of .83 (H), .81 (H without H1), and .61 (C), and .77 (H), .76 (H without H1) and .49 (C) for a more "stringent" definition of violence.

The author claimed that the findings offered support for the HCR-20 in terms of its interrater reliability and validity of the H scale across diverse groups. However, the C scale did not differentiate between groups. The author commented that this is not surprising given that it is meant to measure current dynamic factors, and the outcome measures in this study were all in the past.

Project and Scholarly Work

Douglas, K. S., & Belfrage, H. (2001). Use of the HCR-20 in violence risk management: Implementation and clinical practice. In K. S. Douglas, C. D. Webster, D. Eaves, S. D. Hart, & J. R. P. Ogloff (Eds.), *The HCR-20 Violence Risk Management Companion Manual* (pp. 41-58). Vancouver, BC/Tampa/FL: Mental Health, Law, and Policy Institute, Simon Fraser University/Department of Mental Health Law & Policy, University of South Florida.

Summary

This study was a re-analysis of data that exist from three samples (two cited in this bibliography) for the specific purpose of assessing the degree of change in the Clinical and Risk Management scale and item scores across time and repeated assessments/codings. There were two forensic samples (Belfrage, unpublished raw data; Douglas et al., 1998) and one civil psychiatric sample (Douglas et al., 1999).

In Sample 1 ($n = 193$ civil psychiatric patients), it was possible to compare C scale and item scores at admission and discharge. Each item declined significantly, and the total score declined from 7.21 to 4.05. All drops in scores were large, as assessed by Cohen's d (ds ranged from .89 to 1.75). At admission, 48% of the sample scored in the 8 to 10 range; at discharge, only 3% did so. In Sample 2 ($n = 175$ forensic patients), all C and R items declined, although the drops were not as large as in Sample 1. For the C Scale total score, Cohen's d (.36) indicated a smallish size drop, and for the R Scale, a moderate sized drop (.50). In Sample 3, the C scale did not decline, but the R scale did so moderately ($d = .44$). It is possible that in Sample 3, being drawn from a Swedish forensic facility, that the patients were not as acutely disturbed upon admission (in Sweden

there is no such concept as “Not Guilty by Reason of Insanity” and people are “sentenced” to treatment in the hospital somewhat liberally), and hence change was not observed.

These findings support the conceptualization of the C and R Scales as dynamic (changeable), and hence as appropriate targets for risk management and violence reduction interventions. The fact that the scores changed without direct efforts to change specific HCR-20 factors suggests that declines may be greater with intervention strategies tailored to dynamic HCR-20 risk factors.

See also

Douglas, K. S. (1999). HCR-20 violence risk assessment scheme: “International validity” in diverse settings. In J. Monahan (Chair), *Violence risk assessment: Scientist-practitioner approaches in diverse settings*. Symposium presented at the International joint conference of the American Psychology-Law Society and the European Association of Psychology and Law, July, Dublin, Ireland.

Project and Scholarly Work

Doyle, M. & Dolan, M. (2003, April). *The validity of violence risk assessment instruments in predicting community violence in patients with mental disorder*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Miami Beach, FL.

Summary

This study used a prospective design to examine factors that predicted community violence in discharged mental patients. They also compared the contribution of relatively stable risk factors measured at a baseline period with more dynamic factors measured at different time points. The study used 129 discharged patients (75 of whom were male and 37 were female) and collected data at eight and 24 weeks after discharge. 34 of the patients were forensic cases and 78 were non-forensic. For this study, the authors created three violence categories: physical violence (level 1), any other violence (level 2) and any violence (any level 1 or 2 violence). Level 1 types of violence consisted of: hitting with fist, beating someone up, physically forcing sex on someone, threatening with weapon in hand, using a knife or firing a gun, any violence which results in injury. The authors used three types of factors to assess risk. They were: static factors (PCL-SV, H-scale of HCR-20, VRAG, VRS, MAST, age and DAST), dynamic trait factors (PBQ, NAS, BIS, interpersonal CIRCLE subscales), and dynamic state factors (BPRSS, GAF, BVC, HRS, WARS, and Psy-

chotic rating scales for TCO symptoms and hallucinations). Significant *AUCs* for the static factors regarding any violence were: PCL-SV interpersonal (.64; $p < .05$), PCL-SV social deviance (.66; $p < .01$), PCL-SV total (.67; $p < .01$), VRAG total (.63; $p < .05$), HCR-20 H-scale (.62; $p < .05$). Significant *AUCs* for the static factors regarding level 1 violence were: PCL-SV interpersonal (.68; $p < .05$), PCL-SV social deviance (.67; $p < .05$), PCL-SV total (.69; $p < .01$), VRAG total (.66; $p < .05$), and HCR-20 H-scale (.68; $p < .05$). Significant *AUCs* for the dynamic state factors regarding any violence were: BPRS total (.61; $p < .05$), WARS (.62; $p < .05$). Significant *AUCs* for the dynamic state factors regarding level 1 violence were: BPRS total (.67; $p < .05$) and BPRS hostility-suspicion (.72; $p < .01$).

The scores on the HCR-20 changed significantly across time periods ($p < .001$). The predictive validity for the HCR-20 regarding any violence was: baseline (.65; $p < .01$), discharge (.80; $p < .001$) and eight week follow-up (.69; $p < .01$). The predictive validity for the HCR-20 regarding level 1 violence was: baseline (.63; *ns*), discharge (.80; $p < .001$) and eight week follow-up (.73; $p < .01$).

Regression equations for predicting any violence showed that without the HCR-20 C and R scales included in the model, 72.3% of patients were correctly classified, but with the C and R Scales included in the model, the number correctly classified increased to 85.7%. However, regression equations for predicting level 1 violence showed that without the C and R scales (85.7% correct prediction) was no different than having them included in the model (84.8% correct prediction).

Project and Scholarly Work

Folino, J. O., Marengo, C. M., Marchiano, & Ascazibar, M. (2004). The risk assessment program and the Court of Penal Execution in the province of Buenos Aires, Argentina. *International Journal of Offender Therapy and Comparative Criminology*, 48, 49-58.

Summary

This study provides descriptive data on a risk assessment pilot program that was implemented in September 2001 concurrently with the establishment of the Court of Penal Execution (a special jurisdiction court in Buenos Aires). The sample is a subset of the 1,370 cases admitted to the Court during the first 18 months (September 2001 to February 2002). Of these cases, 105 were serving a sentence either in prison or a forensic psychiatric unit. The sample comprises all cases who were candidates for conditional release ($N = 65$). Participants were men with a mean age of 27.7 years ($SD = 7.8$). There were 55 (85%) convicted offenders and 10 (15%) insanity acquitees. The majority (61.5%) reported a

history of severe drug abuse. A major mental illness was diagnosed in 6% of the sample. Of the participants who were convicted offenders, the evaluation occurred, on average, 1.6 years before their prison term was completed.

The HCR-20, PCL-R, and VRAG were completed using an extensive information gathering process that included criminal records and court files, diagnostic interviews with respondents, interviews with family members, and participant interviews with a clinical practitioner when deemed necessary. The person(s) who completed the assessments was not reported. The mean HCR-20 total score was 18.58 ($SD = 7.63$; range 2-34). The mean PCL-R total score was 20.57 ($SD = 9.05$; range 1-37). The mean VRAG score was 12.17 ($SD = 10.87$; range -13-37).

No significant correlations between any risk measure and judicial resolutions were noted.

Project and Scholarly Work

Hodgins, S, Tengström, A., Östermann, R., Eaves, D., Hart, S. D. Kronstrand, R., Levander, S., Müller-Isberner, R., Tiihonen, J., Webster, C. D., Eronen, M., Freese, R., Jöckel, D., Kreuzer, A., Levin, A., Maas, S., Repo, E., Ross, D., Tuninger, E., Kotilainen, I., Väänänen, K., Vartianen, H., & Vokkolainen, A. (in press). An international comparison of community treatment programs for mentally ill persons. *Criminal Justice and Behavior*.

Summary

This study presents a descriptive overview of a multi-site, international, prospective study concerned primarily with the community aftercare treatment of forensic and civil psychiatric patients. One of the stated goals of the project is to attempt “to validate the HCR-20 in four different cultures,” namely, Canadian, Swedish, German, and Norwegian. Included in this goal is the evaluation of whether there are subtypes of patients for whom the HCR-20 is less accurate, and hence might need revision. The study involves eight data collection sites, two each (one civil psychiatric, one forensic psychiatric) from the four countries.

Reliability analyses to date have included the ratings of four patients each by seven clinicians from different countries. The ICC_1 values were as follows: HCR-20 Total Score (.90); H Scale Score (.94); C Scale Score (.89); R Scale Score (.68). The authors stated that “ICCs are generally very high, indicating excellent inter-rater reliability.”

At present, the project is not far enough along to provide predictive validity analyses. However, criterion-related validity has been partially evaluated through correlating the seven clinicians’ ratings with those made by the authors of the HCR-20 on the four cases. The ICC_1 values for

these analyses were as follows: HCR-20 Total Scale Score (.99); H Scale Score (.85); C Scale Score (.99); R Scale Score (.96).

Future research reports from this study will provide information on predictive validity in both forensic and civil psychiatric patients, from four countries, on ratings made prospectively through both clinical interview and file review methodology.

Project and Scholarly Work

Freese, R., Hiscoke, U. & Hodgins, S. (May, 2002). *The treatment of mentally ill patients with a history of criminality or violence: What works and what doesn't work?* Paper presentation at 11th congress of the association of European psychiatrists, Stockholm, Sweden.

Summary

These authors used a prospective design to determine if any markers at time of discharge (such as PCL-R or HCR-20 scores) would be predictive of future violent acts. They used a male only sample with schizophrenia spectrum disorder and followed 128 individuals up to 6 months after discharge from either forensic (aggressive individuals) or general psychiatry (non-aggressive individuals) clinics and then they were able to follow up 78 of the original 128 for a second six month period to make it a 1 year follow up for this smaller group. This report was based on the Hodgins et al (in press) sample reported above.

At discharge, the aggressive individuals showed a mean of 15.1 ($SD = 8.1$) on the PCL-R and a mean of 22.2 ($SD = 6.5$) on the HCR-20 score. The HCR-20 subscale means were: H-scale mean 12.8 ($SD = 3.7$), C-scale mean 3.9 ($SD = 2.8$) and R-scale mean 5.6 ($SD = 2.2$). At discharge, the non-aggressive individuals showed a mean of 12.2 ($SD = 7.2$) on the PCL-R and a mean of 17.2 ($SD = 6.5$) on the HCR-20 score. The HCR-20 subscale means were: H-scale mean 10.4 ($SD = 4.4$), C-scale mean 3.1 ($SD = 1.9$) and R-scale mean 3.8 ($SD = 2.2$).

Results showed that the PCL-R did not predict aggressive behavior. Results also showed that the HCR-20 H-scale did not predict future violent acts. However, the C and R scales did predict future aggressive behavior. The strongest predictors of future violent behavior in this sample were increases in anxiety and depressive symptoms over time.

Logistic regression analyses to predict aggressive behaviour during the first follow-up period indicated the HCR-20 total score increased risk by 1.1 times per unit and the R-scale increased risk by 1.5 times per unit (TCO and a score of 5+ on PANSS positive symptoms increased risk

by 1.2 and 5.2 times, respectively). TCO and HCR-20 indices did not remain significant when PCL-R total score and diagnosis of substance abuse/dependence were controlled for. For the second follow-up period, odds ratios for HCR-20 total, C-scale, and R-scale were 1.2, 2.1, and 2.2, respectively. These values are substantially lower in comparison to values for the other measures (e.g., a score of 5+ on PANSS positive symptoms yielded an odds ratio of 34.0). As was the case for the first follow-up period, no HCR-20 indices remained significant once PCL-R total score and diagnosis of substance abuse/dependence were controlled for.

Project and Scholarly Work

Tengström, A. & Hodgins, S. (in press). Criminal behavior of forensic and general psychiatric patients with schizophrenia: Are they different? *Acta Psychiatrica Scandinavica*.

Summary

This study compared the rate of past criminal behavior among male patients being discharged from forensic and general psychiatric hospitals in four countries (same study as Hodgins et al., in press, above). Clinicians also assessed the risk of violent behavior in the future using the Psychopathy Checklist and the HCR-20. The sample consisted of 110 forensic patients and 47 general psychiatry patients. Patients, all of whom were men, had either schizophrenia, schizoaffective, or schizophreniform disorder. Within each site, each patient being discharged from a forensic hospital was matched to a patient of the same sex, age, and primary diagnosis being discharged from a general psychiatric service.

A number of comparisons between the forensic and general psychiatric groups was undertaken (e.g., age at discharge, various criminal history indices, type and history of psychiatric admissions). In addition to criminal history variables, comparisons of a more clinical nature were made between the general psychiatric patients with and without criminal histories. Relative to general psychiatric patients without a criminal history, those with a criminal history had significantly higher HCR-20 total scores ($X = 21.00$, $SD = 5.54$ vs. $X = 15.54$, $SD = 6.18$; $t(47) 2.70$, $p = .01$), R-scale scores ($X = 6.25$, $SD = 1.42$ vs. $X = 4.23$, $SD = 2.21$; $t(47) 3.64$, $p = .001$), and PCL-R total scores ($X = 13.27$, $SD = 6.36$ vs. $X = 9.34$, $SD = 5.42$; $t(47) 2.07$, $p = .04$). No significant difference between patients with versus without a criminal history was observed for H-scale scores ($X = 9.75$, $SD = 5.05$ vs. $X = 7.29$, $SD = 3.16$; $t(47) 1.59$, $p = .13$), C-scale scores ($X = 5.00$, $SD = 1.81$ vs. $X = 4.03$, $SD = 1.98$; $t(47) 1.50$, $p = .14$), PCL-R Factor 1 scores ($X = 4.67$, $SD = 3.42$ vs. $X = 3.43$, $SD = 2.62$; $t(47) 1.31$, $p = .19$), or PCL-R Factor 2 scores ($X = 7.41$, $SD = 2.67$ vs. $X = 5.60$, $SD = 3.67$; $t(47) 1.56$, $p =$

.13). A global clinical judgment (it was not specified whether this judgment was based solely on information gathered during the completion of the HCR-20) of risk for future behavior (low, moderate, high) over the subsequent 6 months did not distinguish the two groups. It was not specified whether this clinical judgment was made according to the SPJ model, or was unstructured.

The patients with a criminal history were assessed as having a greater risk for violent behavior in the community after discharge as indicated by higher total scores on the PCL-R and on the HCR-20. Lastly, the global clinical judgment of risk of future violence did not distinguish between the two study groups (forensic and general psychiatry).

END OF MIXED SETTING STUDIES

JUVENILE OFFENDER AND FORENSIC SETTINGS

Project and Scholarly Work

MacEachern, A. (2001). *Violence risk assessment: Comparing the predictive validity of the HCR-20 and the SAVRY in a population of adolescent offenders*. Unpublished Master's Thesis, Department of Psychology, Simon Fraser University.

Summary

This study compared the HCR-20 and the SAVRY in a sample of 108 male juvenile offenders who were referred from court for inpatient psychiatric assessment (36 randomly selected who were nonrecidivists, 36 nonviolent recidivists, and 36 violent recidivists). This was a pseudo-prospective/retrospective follow-up study conducted from comprehensive youth justice, police, mental health, medical, and social-demographic files. The juvenile offenders were, on average, 15.3 at admission, and 25.1 at follow-up. Hence, this study evaluates the predictive validity of later adult violence of adolescent offenders. Follow-up national criminal records were used to code violence. Most participants were Caucasian.

Interrater reliability, based on a subset of 36 files, was good. ICC_1 for Total, H, C, and R scores was .86, .88, .80, and .77. Pearson correlation coefficients between the HCR-20 Total Score, H, C, and R and violence were as follows: .46, .42, .35, and .44.. Corresponding AUC values were .79, .76, .73, and .78.

Mean HCR-20 scores differed significantly between non-recidivists, nonviolent recidivists, and violent recidivists. Total, H, C, and R scale scores across these three groups were as follows: 15.1 vs. 20.0 vs. 23.9; 6.1 vs. 7.8 vs. 9.8; 4.5 vs. 5.8 vs. 6.5; 4.5 vs. 6.5 vs. 7.6.

Comparison to the SAVRY showed that the HCR-20 produced somewhat larger correlations and AUC values, although there were no tests of significance between the two measures. Logistic regression analyses with all HCR-20 and SAVRY subscales showed that the HCR-20 Total, H, C, and R scale scores remained significant in the final model along with the SAVRY total score.

END OF JUVENILE SETTINGS

REFERENCES

HCR-20 Professional Manuals

Douglas, K. S., Webster, C. D., Hart, S. D., Eaves, D., & Ogloff, J. R. P. (Eds.) (2001). *HCR-20: Violence risk management companion guide*. Vancouver, BC/Tampa, FL: Mental Health, Law, and Policy Institute, Simon Fraser University/Department of Mental Health Law & Policy, University of South Florida.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997b). *HCR-20: Assessing the Risk for Violence (Version 2)*. Vancouver: Mental Health, Law, and Policy Institute, Simon Fraser University.

Webster, C.D., Eaves, D., Douglas, K.S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Vancouver, B.C.: Mental Health Law and Policy Institute, and Forensic Psychiatric Services Commission of British Columbia.

Scholarly works with HCR-20 data that were summarized in the bibliography

Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: Integrating research and clinical practice. *Journal of Forensic Psychiatry*, 9, 328-338.

Belfrage, H. & Douglas, K. (2002). Treatment effects on forensic psychiatric patients measured with the HCR-20 violence risk assessment scheme. *International Journal of Forensic Mental Health*, 1, 25-36.

Belfrage, H., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: A prospective study in two maximum-security correctional institutions. *Journal of Forensic Psychiatry*, 11, 167-175.

Belfrage, H. Fransson, G., & Strand, S. (2004). Management of violent behavior in the correctional system using qualified risk assessments. *Legal and Criminological Psychology*, 9, 11-22.

Brown, L. K. (2001). *Assessing risk for elopement and breaches of conditional release in insanity acquittees*. Unpublished Ph.D. Dissertation. Simon Fraser University, Burnaby, British Columbia, Canada.

Claix, A., Pham, T., & Willocq, L. (2002, March). *Evaluation of the HCR-20 (Historical-Clinical-Risk management) in a Belgian forensic population*. Poster presented

at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.

Cooke, D. J., Michie, C., & Ryan, J. (2001). *Evaluating risk for violence: A preliminary study of the HCR-20, PCL-R and VRAG in a Scottish Prison Sample*. Report prepared for the Scottish Prison Service.

Côté, G. (2001, April). *Violent behaviour, PCL-R and HCR-20 among involuntary inpatients, forensic patients and severely mentally disordered inmates*. Paper presented at the First Annual Meeting of the International Association of Forensic Mental Health Services, Vancouver, BC.

de Vogel, V. & de Ruiter, C. (2004). Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients. *The Journal of Forensic Psychiatry and Psychology*, 15, 145-164.

de Vogel, V., & de Ruiter, C. (under review). *Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20*. Manuscript submitted for publication

de Vogel, V. & de Ruiter, C. (in press). The HCR-20 in personality disordered female offenders: A comparison with a matched sample of males. *Clinical Psychology and Psychotherapy*.

de Vogel, V., de Ruiter, C. Hildebrand, M., Bos, B. & van de Ven, P. (2004). Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health*, 3, 149-165.

de Vogel, V., de Ruiter, C., & Vandeputte, C. (2001, November). *Implementation of the HCR-20 and SVR-20 in a Dutch forensic psychiatric hospital*. Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden.

Dernevik, M. (1998). Preliminary findings on reliability and validity of the Historical-Clinical-Risk Assessment in a forensic psychiatric setting. *Psychology, Crime, and Law*, 4, 127-137.

Dernevik, M., Falkheim, M., Holmqvist, R. & Sandell, R. (2001) Implementing Risk Assessment: Clinical Judgment Revisited. In D. Farrington, C. Hollin, & M. McMurrin (Eds.). *Sex and Violence: The Psychology of Crime and Risk Assessment*, Harwood Academic Press. London.

- Dernevik, M., Grann, M., & Johansson, S. (2002). Violent behaviour in forensic psychiatric patients: Risk assessment and different risk management levels using the HCR-20. *Psychology, Crime, and Law*, 8, 83-111.
- Dolan, M., & Khawaja, A. (2004). The HCR-20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior*, 30, 469-483.
- Douglas, K. S., & Belfrage, H. (2001). Use of the HCR-20 in violence risk management: Implementation and clinical practice. In K. S. Douglas, C. D. Webster, D. Eaves, S. D. Hart, & J. R. P. Ogloff (Eds.), *The HCR-20 Violence Risk Management Companion Manual* (pp. 41-58). Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University, and Department of Mental Health Law & Policy, University of South Florida.
- Douglas, K. S., Klassen, C., Ross, C., Hart, S. D., & Webster, C. D. (1998, August). *Psychometric properties of HCR-20 violence risk assessment scheme in insanity acquittees*. Paper presented at the Annual meeting of the American Psychological Association, San Francisco, CA.
- Douglas, K. S., & Ogloff, J. R. P. (2003). Multiple facets of risk for violence: The impact of judgmental specificity on structured decisions about violence risk. *International Journal of Forensic Mental Health*, 1, 19-34.
- Douglas, K. S., & Ogloff, J. R. P. (2003). The impact of confidence on the accuracy of structured professional and actuarial violence risk judgments in a sample of forensic psychiatric patients. *Law and Human Behavior*, 27, 573-587.
- Douglas, K. S., Ogloff, J. R. P., & Hart, S. D. (2003). *Evaluation of a model of violence risk assessment among forensic psychiatric patients*. *Psychiatric Services*, 54, 1372-1379.
- Douglas, K. S., Ogloff, J. R. P., & Nicholls, T. L. (1997, August). *Violence by psychiatric patients: Validity of the HCR-20 Scheme and the Psychopathy Checklist: Screening Version*. Paper presented at the annual convention of the American Psychological Association, Chicago.
- Douglas, K. S., Ogloff, J. R. P., & Nicholls, T. L. (1997, June). *The role of personality disorders in community violence among civil psychiatric patients*. Paper presented at the Fifth International Congress of the Disorders of Personality, Vancouver, B.C.
- Douglas, K. S., Ogloff, J. R. P., & Nicholls, T. L. (1997, June). *Assessing the risk for inpatient psychiatric violence*. Paper presented at the annual convention of the Canadian Psychological Association, Toronto.
- Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology*, 67, 917-930.
- Douglas, K. S., & Webster, C. D. (1999). The HCR-20 violence risk assessment scheme: Concurrent validity in a sample of incarcerated offenders. *Criminal Justice and Behavior*, 26, 3-19.
- Douglas, K.S., Webster, C.D., & Wintrup, A. (1996, August). *The HCR-20 Risk Assessment Scheme: Psychometric properties in two samples*. Poster presented at the annual convention of the American Psychological Association, August, 1996, Toronto, Ontario.
- Douglas, K. S., Yeomans, M. & Boer, D. (in press). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior*.
- Doyle, M. & Dolan, M. (2003, April). *The validity of violence risk assessment instruments in predicting community violence in patients with mental disorder*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Miami Beach, Florida.
- Doyle, M., Dolan, M., & Mc Govern, J. (2002). The validity of North American risk assessment tools in predicting inpatient violent behaviour in England. *Legal and Criminological Psychology*, 7, 141-154.
- Dunbar, E. (2003). Symbolic, relational, and ideographical signifiers of bias-motivated offenders: Toward a strategy of assessment. *American Journal of Orthopsychiatry*, 73, 203-211.
- Folino, J. O., Marengo, C. M., Marchiano, & Ascazibar, M. (2004). The risk assessment program and the Court of Penal Execution in the province of Buenos Aires, Argentina. *International Journal of Offender Therapy and Comparative Criminology*, 48, 49-58.
- Freese, R., Hiscoke, U. & Hodgins, S. (2002, May). *The treatment of mentally ill patients with a history of criminality or violence: What works and what doesn't work?* Paper presentation at 11th congress of the association of European psychiatrists, Stockholm, Sweden.

- Fujii, D., Lichton, A., & Tokioka, A. (2004). Structured professional judgment versus actuarial data in violence risk prediction using the Historical Clinical Risk Management-20. Manuscript under review.
- Fujii, D., Tokioka, A., & Lichton, A. (in press). Cultural differences in violence risk prediction of psychiatric inpatients using the Historical Clinical Risk Management-20. *Psychiatric Services*.
- Grann, M., Belfrage, H., & Tengström, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behavior*, 27, 97-114.
- Grann, M., & Långström, N. (in press). Actuarial assessment of risk for violence: To weigh or not to weigh? *Criminal Justice and Behavior*.
- Gray, N. S., Hill, C., McGleish, A., Timmons, D., MacCulloch, M. J., & Snowden, R. J. (2003). Prediction of violence and self-harm in mentally disordered offenders: A prospective study of the efficacy of the HCR-20, PCL-R, and psychiatric symptomatology. *Journal of Consulting & Clinical Psychology*, 71, 443-451.
- Gray, N. S., Snowden, R. J., MacCulloch, S., Phillips, H., Taylor, J. & MacCulloch, M. J. (2004). Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally disordered offenders: A comparative study of HCR-20, PCL:SV, and OGRS. *Journal of Consulting and Clinical Psychology*, 72, 523-530.
- Gretenkord, L., Muller-Isberner, R., Ozokoyay, K. & Sommer, J. (2002, March). *Validating the HCR-20: relationship between levels of security and the CR-10 score in hospital order treatment*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.
- Grevatt, M., Thomas-Peter, B., & Hughes, G. (2004). Violence, mental disorder and risk assessment: Can structured clinical assessments predict the short-term risk of inpatient violence? *The Journal of Forensic Psychiatry & Psychology*, 15, 278-292.
- Hilterman, E. & Chakhssi, F. (2002, March). *Prospective assessment of risk: comparing HCR-20, Behavioural Status Index (BSI) and Leave Risk Assessment (LRA)*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.
- Hodgins, S, Tengström, A., Östermann, R., Eaves, D., Hart, S. D. Kronstrand, R., Levander, S., Müller-Isberner, R., Tiihonen, J., Webster, C. D., Eronen, M., Freese, R., Jöckel, D., Kreuzer, A., Levin, A., Maas, S., Repo, E., Ross, D., Tuninger, E., Kotilainen, I., Väänänen, K., Vartiainen, H., & Vokkolainen, A. (in press). An international comparison of community treatment programs for mentally ill persons. *Criminal Justice and Behavior*.
- Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, 28, 471-489.
- Logan, C., Doyle, M. & Dolan, M. (2001, November). *Risk assessment in English male offenders who have committed single acts of serious violence or sexual violence: how can the past be used to predict the future?* Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden.
- Logan, C. & Watt, K. (2001, November). *Structured professional guidelines approaches to risk assessment: Single practitioner vs. multidisciplinary team administration*. Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden.
- MacEachern, A. (2001). *Violence risk assessment: Comparing the predictive validity of the HCR-20 and the SAVRY in a population of adolescent offenders*. Unpublished Master's Thesis, Department of Psychology, Simon Fraser University.
- McNiel, D., Gregory, A., Lam, J., Binder, R., & Sullivan, G. (2003). Utility of decision support tools for assessing acute risk of violence. *Journal of Consulting and Clinical Psychology*, 71, 945-953.
- Müller-Isberner, R., & Jockel, D. (1997, September). *The implementation of the HCR-20 in a German hospital order institution*. Paper presented at the Seventh European Conference on Psychology and Law, Solna, Sweden.
- Müller-Isberner, R., Sommer, J., Özokoyay, K., & Freese, R. (1999, November). *Clinical use of the HCR-20 for predicting violence in a German Forensic Psychiatric Hospital*. Paper presented at the International Conference on Risk Assessment and Management: Implications for Prevention of Violence, Vancouver, BC, Canada.
- Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (1997, August). *Comparing risk assessments with female and male psychiatric outpatients: Utility of the HCR-20 and*

- Psychopathy Checklist: Screening Version*. Paper presented at the annual convention of the American Psychological Association, Chicago.
- Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (1997, June). *Comparing risk assessments with female and male psychiatric inpatients: Utility of the HCR-20 and Psychopathy Checklist: Screening Version*. Paper presented at the annual convention of the Canadian Psychological Association, Toronto.
- Nicholls, T. N., Ogloff, J. R. P., & Douglas, K. S. (2004). Assessing risk for violence among male and female psychiatric patients: The HCR-20, PCL:SV, and VSC. *Behavioral Sciences and the Law*, 22, 127-158.
- Nicholls, T. L., Vincent, G. M., Whitemore, K. E., & Ogloff, J. R. P. (1999, November). *Assessing risk of inpatient violence in a sample of forensic psychiatric patients: Comparing the PCL:SV, HCR-20, and VRAG*. Paper presented at the International Conference on Risk Assessment and Risk Management, Vancouver, BC.
- Ogloff, J. R. P., Douglas, K. S., Nicholls, T. N., & Grant, I. (1997, November). *Civil commitment and risk for violence in psychiatric patients*. Paper presented at the annual meeting of the Pinel Institute, Montreal, Quebec, Canada.
- Ogloff, J. R. P., Douglas, K. S., Nicholls, T. N., & Grant, I. (1997, May). *Involuntary civil commitment: Risk assessment, sex differences, and review panel decision making*. Paper presented at the Annual Convention of the Law and Society Association, St. Louis, Missouri.
- Pham, T. (2001, November). *Assessing risk for violence among Belgian offenders*. Paper presented at the International Conference, Violence Risk Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden.
- Pham, T. H., Claix, A., & Remy, S. (2000, June). *Assessment of the HCR-20 in a Belgian prison sample*. Paper presented at the 4th European Congress on Personality Disorders. Paris, France.
- Philipse, M. (2002, March). *Post-dictive validity of the HCR-20 in a Dutch forensic psychiatric sample*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany.
- Philipse, M., Erven, T. van, & Peters, J. (2002). Risicotaxatie in de tbs: van geloof naar empirie. [Risk assessment in tbs: from belief to empiricism.] *Justitiële Verkenningen [Judicial Explorations]*, 28(8), 77-93.
- Ross, D. J., Hart, S. D., Eaves, D., & Webster, C. D. (1999, November). *The relationship between the HCR-20 and BC Review Board decisions on the release of forensic psychiatric inpatients*. Paper presented at the International Conference on Risk Assessment and Risk Management, Vancouver, BC.
- Ross, D. J., Hart, S. D., Eaves, D., & Webster, C. D. (2001, April). *The relationship between the HCR-20 and community violence in a sample of NCRMD outpatients*. Paper presented at the Founding Conference of the International Association of Forensic Mental Health Services, Vancouver, BC.
- Ross, D. J., Hart, S. D., & Webster, C. D. (1998). *Aggression in psychiatric patients: Using the HCR-20 to assess risk for violence in hospital and in the community*. Unpublished manuscript.
- Scharin, C. (1999). *Bedömning av återfallsrisk hos rättspsykiatriskt undersökta personer: En utvärdering av skattningsskalan HCR-20*. Unpublished manuscript.
- Strand, S., & Belfrage, H. (2001). Comparison of HCR-20 scores in violent mentally disordered men and women: Gender differences and similarities. *Psychology, Crime and Law*, 7, 71-79.
- Strand, S., Belfrage, H., Fransson, G., & Levander, S. (1999). Clinical and risk management factors in risk prediction of mentally disordered offenders: More important that actuarial data? *Legal and Criminological Psychology*, 4, 67-76.
- Tengström, A. & Hodgins, S. (in press). Criminal behavior of forensic and general psychiatric patients with schizophrenia: Are they different? *Acta Psychiatrica Scandinavica*.
- Tengström, A. (2001). Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry*, 55, 243-249.
- Urheim, R., Jakobsen, D., & Rasmussen, K. (2003, August). *Dimensions of inpatient aggressive behavior in a security ward: What is being "predicted"?* Paper presented at the 5th Nordic Symposium on Forensic Psychiatry, Ystad, Sweden.
- Vincent, G. M. (1998). *Criminal responsibility after Bill C-30: Factors predicting acquittal and lengths of confinement in British Columbia*. Unpublished master's thesis, Simon Fraser University, Burnaby, British Columbia, Canada.

- Vincent, G. M., Ross, D. J., Whittemore, K., Eaves, D., Hart, S. D., Ogloff, J. R. P., & Webster, C. D. (2001, April). *Using the HCR-20: File-based researcher ratings vs. file + interview-based clinician ratings*. Paper presented at the Founding conference of the International Association of Forensic Mental Health Services, Vancouver, BC.
- Watt, A., Topping-Morris, B., Rogers, P., Doyle, M., & Mason, T. (2003). Pre-admission nursing assessment in a Welsh medium secure unit (1991-2000): Part 2 – comparison of traditional nursing assessment with the HCR-20 risk assessment tool. *International Journal of Nursing Studies, 40*, 657-662.
- Webster, C. D., Douglas, K. S., Belfrage, H., & Link, B. (2000). Capturing change: An approach to managing violence and improving mental health. In S. Hodgins & R. Müller-Isberner (Eds.), *Violence among the mentally ill* (pp. 119-144). Dordrecht, Netherlands: Kluwer/Academic.
- Whittemore, K. E. (1999). *Releasing the mentally disordered offender: Disposition decisions for individuals found unfit to stand trial and not criminally responsible*. Unpublished Ph.D. Dissertation, Simon Fraser University, Burnaby, BC.
- Wintrup, A. (1996). *Assessing risk of violence in mentally disordered offenders with the HCR-20*. Unpublished master's thesis, Simon Fraser University, Burnaby, BC.
- McNeil, D. E., Sandberg, D. A., & Binder, R. L. (1998). The relationship between confidence and accuracy in clinical assessment of psychiatric patients' potential for violence. *Law and Human Behavior, 22*, 655-669.
- Metz, C. E. (1984). Statistical analysis of ROC data in evaluating diagnostic performance. In D. E. Herbert & R. H. Myers (Eds.), *Multiple regression analysis: Applications in the health sciences* (pp. 365-384). Washington, DC: American Institute of Physics.
- Mossman, D. (1994a). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62*, 783-792.
- Mossman, D. (1994b). Further comments on portraying the accuracy of violence predictions. *Law and Human Behavior, 18*, 587-594.
- Mossman, D., & Somoza, E. (1991). ROC curves, test accuracy, and the description of diagnostic tests. *Journal of Neuropsychiatry and Clinical Neurosciences, 3*, 330-333.
- Overall, J. E., & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports, 10*, 799-812.
- Rice, M. E. (1997). Violent offender research and implications for the criminal justice system. *American Psychologist, 52*, 414-423.
- Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*, 737-748.
- Vida, S. (1997). *AccuROC nonparametric receiver operating characteristic analysis* (version 3.1) [Computer software]. Montreal, Quebec, Canada: Accumetric Corporation.
- Wong, S., & Gordon, A. (2001). The Violence Risk Scale. *The Bulletin of the International Society for Research on Aggression, 23*.

Other scholarly works cited in the bibliography

- Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology, 42*, 861-865.
- Copas, J., & Marshall, P. (1998). The offender group reconviction scale: A statistical reconviction score for use by probation officers. *Applied Statistics, 47*, 159-171.
- Hsiao, J. K., Bartko, J. J., & Potter, W. Z. (1989). Diagnosing diagnoses: Receiver operating characteristic methods and psychiatry. *Archives of General Psychiatry, 46*, 664-667.
- Lusted, L. B. (1978). General problems in medical decision making with comments on ROC analysis. *Seminars in Nuclear Medicine, 8*, 299-306.

Related Disseminations/Academic Commentary

Book/Test Reviews

- Arbisi, P. A. (2003). Review of the HCR-20: Assessing Risk for Violence. From B. S. Plake, J. C. Impara, & R.

A. Spies (Eds.), *The fifteenth mental measurements yearbook* [Electronic version]. Retrieved May 7, 2003, from the Buros Institute's Test Reviews Online website: <http://www.unl.edu/buros>.

- Buchanan, A. (2001). [Review of the book, *HCR-20: Assessing risk for violence*, Version 2]. *Criminal Behaviour and Mental Health*, *11*, S77-S89.
- Cooper, C. (2003). Review of the HCR-20: Assessing Risk for Violence. From B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* [Electronic version]. Retrieved May 7, 2003, from the Buros Institute's Test Reviews Online website: <http://www.unl.edu/buros>.
- Mossman, D. (2000). Evaluating violence risk 'by the book': A review of HCR-20: Assessing risk for violence, Version 2 and The manual for the sexual violence risk – 20. *Behavioral Sciences and the Law*, *18*, 781-789.
- Witt, P. H. (2000). A practitioner's view of risk assessment: The HCR-20 and SVR-20. *Behavioral Sciences and the Law*, *18*, 791-798.

Translations

Translated into Dutch as:

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (2000). *HCR-20: Beoordelen van het risico van geweldadig gedrag, Versie 2* (M. Philipse, C. de Ruiter, M. Hildebrand, & Y. Bouman, Eds. and Trans.). Nijmegen/Utrecht, The Netherlands: Prof. Mr. W. P. J. Pompestichting & Dr. Henri van der Hoeven Stichting (Original work published 1997).

Translated into French as:

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1999). *HCR-20: Évaluation du risque de violence, version 2*. Burnaby, British Columbia: Mental Health, Law, & Policy Institute, Simon Fraser University (Original work published 1997).

Translated into German as:

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1998). *Die Vorhersage von Gewalttaten mit dem HCR-20* (R. Müller-Isberner, D. Jöckel, & S. G. Cabeza, Eds. and Trans.). Haina, Germany: Institut für Forensische Psychiatrie (Original work published 1997).

Translated into Swedish as:

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Bedömning av risk för framtida våld. Manual med instruktioner och kommentarer* (H. Belfrage & G. Fransson, Eds. and Trans.). Växjö: Psykiatriskt regionvårdscentrum Landstinget Kronoberg (Original work published 1997).

Related Chapters and Articles by HCR-20 Authors

- Borum, R., & Douglas, K. S. (2003). New directions in violence risk assessment. *Psychiatric Times*, *XX* (3), 102-103.
- Douglas, K.S., Cox, D.N., & Webster, C.D. (1999). Violence risk assessment: Science and practice. *Legal and Criminological Psychology*, *4*, 149-184.
- Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, *29*, 617-658.
- Douglas, K.S., Macfarlane, E., & Webster, C.D. (1996). Predicting dangerousness in the contemporary Canadian mental health and criminal justice systems. *Canada's Mental Health*, *43*, 4-11.
- Douglas, K.S., & Webster, C.D. (1999). Predicting violence in mentally and personality disordered individuals. In R. Roesch, S.D. Hart, & J. R. P. Ogloff (Eds.), *Psychology and law: The state of the discipline* (pp. 175-239). New York: Plenum.
- Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology*, *3*, 121-137.
- Lyon, D. R., Hart, S. D., & Webster, C. D. (2001) Violence and risk assessment. In R. Schuller & J. R. P. Ogloff (Eds.), *Law and Psychology: Canadian Perspectives* (pp. 314-350). Toronto: University of Toronto Press.
- Webster, C.D., Douglas, K.S., Eaves, S.D., & Hart, S.D. (1997a). Assessing risk of violence to others. In C.D. Webster & M.A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 251-277). New York: Guilford.
- Webster, C. D. & Jackson, M. A. (Eds.), *Impulsivity: Theory, assessment, and treatment*. New York: Guilford.
- Borum, R. (1998). Forensic assessment instruments. In G. P. Koocher, J. C. Norcross, & S. S. Hill (Eds.), *Psychologists' desk reference* (pp. 487-491). New York, NY: Oxford University Press.

Published Peer Academic Commentary²

² All of the disseminations in this section spend at least some time on the HCR-20. Some are devoted largely to the HCR-20, whereas others provide more circumscribed discussion.

- sented at Correctional Service of Canada (Pacific Region) Risk Assessment Training Seminar, Abbotsford, British Columbia, March 17, 1997.
- Douglas, K. S. (1997, May). *Violence in mentally disordered offenders and forensic psychiatric patients: The need for systematized assessment procedures*. Paper presented Continuing Education Conference of the British Columbia Provincial Court Judge's Association, Penticton, B. C., May 2, 1997.
- Douglas, K. S. (1998, March). *The HCR-20 Violence Risk Assessment Scheme: A summary of research findings*. Symposium paper presented at the Biennial meeting of the American Psychology-Law Society, March 4 - 6, 1998, Redondo Beach, CA.
- Douglas, K. S. (1999). HCR-20 violence risk assessment scheme: "International validity" in diverse settings. In J. Monahan (Chair), *Violence risk assessment: Scientist-practitioner approaches in diverse settings*. Symposium presented at the International joint conference of the American Psychology-Law Society and the European Association of Psychology and Law, July, Dublin, Ireland.
- Douglas, K.S. Webster, C.D., Eaves, D., Wintrup, A., & Hart, S.D. (1996, March). *A new scheme for the assessment of dangerousness and the prediction of violence*. Paper presented at the biennial conference of the American Psychology-Law Society (division 41 of the American Psychological Association), Hilton Head, South Carolina.
- Grann, M. (2001, April). *Actuarial assessment of violence risk: To weigh or not to weigh?* Paper presented at the International conference of the International Association of Forensic Mental Health Services, Vancouver, BC, Canada.
- Hart, S. D. (1996, March). *Conceptual issues in predicting violence: Some unasked and unanswered questions about risk assessment*. Paper presented at the biennial meeting of the American Psychology-Law Society, Hilton Head, South Carolina.
- Hart, S. D. (1998, March). *Actuarial risk assessment: Unethical and illegal?* Paper presented at the biennial meeting of the American Psychology-Law Society, Redondo Beach, California.
- Kullgren, G. (2001, April). *Clinical (C5) and risk (R-5) scores according to the HCR-20 as related to treatment factors and criminal recidivism*. Paper presented at the International conference of the International Association of Forensic Mental Health Services, Vancouver, BC, Canada.
- Ross, D. (2001, April). *The relationship between HCR-20 scores and community violence in a sample of NCRMD outpatients*. Paper presented at the International conference of the International Association of Forensic Mental Health Services, Vancouver, BC, Canada.
- Vincent, G. (2001, April). *Using the HCR-20: File based researcher ratings vs. file and interview-based psychiatrist ratings*. Paper presented at the International conference of the International Association of Forensic Mental Health Services, Vancouver, BC, Canada.

Related Professional Manuals

- Augimeri, L. K., Koegl, C. J., Webster, C. D., & Levene, K. S. (2001). *Early assessment risk list for boys (EARL-20B): Version 2*. Toronto: Earls court Child and Family Centre.
- Augimeri, L. K., Webster, C. D., Koegl, C. J., & Levene, K. S. (1998). *Early assessment risk list for boys (EARL-20B): Consultation Edition*. Toronto: Earls court Child and Family Centre.
- Borum, R., Bartel, P., & Forth, A. (2002). *SAVRY: Manual for the structured assessment of violence risk in youth (Version 1: Consultation Edition)*. Tampa, FL: Florida Mental Health Institute, University of South Florida.
- Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1998). *Manual for the Sexual Violence Risk — 20*. Vancouver: British Columbia Institute Against Family Violence.
- Eaves, D., Douglas, K. S., Webster, C. D., Ogloff, J. R. P., & Hart, S. D. (2000). *Assessing dangerous and long-term offenders: A guide for mental health professionals*. Vancouver: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Hare, R. D. (1991). *Manual for the Hare Psychopathy Checklist - Revised*. Toronto: Multi-Health Systems.
- Hart, S. D., Cox, D., & Hare, R. D. (1995). *The Hare psychopathy checklist: Screening version (PCL:SV)*. Toronto, Ontario: Multi-Health Systems.
- Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The risk for sexual violence protocol*. Burnaby, BC: Simon Fraser University.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1999). *Manual for the Spousal Assault Risk Assessment Guide (version 3)*. Vancouver: British Columbia Institute on Family Violence.
- Roesch, R., Zapf, P. A., Eaves, D., & Webster, C. D. (1998). *Fitness Interview Test (Revised Edition)*. Vancouver, British Columbia: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The violence prediction scheme:*

Assessing dangerousness in high risk men. Toronto:
University of Toronto: Centre of Criminology.

An Invitation and Request

We invite any and all qualified persons to carry out research on the HCR-20. Our hope is to understand its psychometric properties as well as possible, and this requires empirical investigation at different sites and in various diverse samples.

We ask any persons who may have carried out research on the HCR-20 to please forward a copy of any presentations, manuscripts, or publications that emerge therefrom. Documents may be sent to the address below. We aim to keep this annotated bibliography as current as possible, with as much existing research as possible.

Thank You!

<h2>Contact information</h2>

To contact Kevin Douglas at Simon Fraser University:

Kevin S. Douglas, LL.B., Ph.D.
Assistant Professor
Department of Psychology
Simon Fraser University
8888 University Drive
Burnaby, British Columbia
Canada, V5A 1S6
Phone: 604-268-7050
Fax: 604-291-3427
Email: douglask@sfu.ca

To contact Simon Fraser University

Mental Health, Law, and Policy Institute
Simon Fraser University
8888 University Drive
Burnaby, British Columbia
Canada, V5A 1S6
Phone: 604-291-5868
Fax: 604-291-3427
Email: mhlpi@sfu.ca
Internet: <http://www.sfu.ca/psychology/groups/mhlpi>

To contact Psychological Assessment Resources, Inc.

<http://www.parinc.com>