

Violence risk prediction

Clinical and actuarial measures and the role of the Psychopathy Checklist

M. DOLAN and M. DOYLE

Background Violence risk prediction is a priority issue for clinicians working with mentally disordered offenders.

Aims To review the current status of violence risk prediction research.

Method Literature search (Medline). Key words: violence, risk prediction, mental disorder.

Results Systematic/structured risk assessment approaches may enhance the accuracy of clinical prediction of violent outcomes. Data on the predictive validity of available clinical risk assessment tools are based largely on American and North American studies and further validation is required in British samples. The Psychopathy Checklist appears to be a key predictor of violent recidivism in a variety of settings.

Conclusions Violence risk prediction is an inexact science and as such will continue to provoke debate. Clinicians clearly need to be able to demonstrate the rationale behind their decisions on violence risk and much can be learned from recent developments in research on violence risk prediction.

Declaration of interest None.

Violence risk assessment and management are key components of clinical practice (Monahan, 1992). In the UK, however, the adequacy and accuracy of risk prediction has been questioned in several inquiries into serious incidents involving mentally disordered patients (Reed, 1997), and there is a growing emphasis on systematising protocols for risk assessment and management. More recently, the Government highlighted the issue of risk in those with 'dangerous severe personality disorders' (DSPD) and outlined new proposals for dealing with this challenging but ill-defined group. Risk prediction is once again high on the public, political and clinical agenda. Here we review progress on violence risk assessment (particularly, violent recidivism) in the mentally disordered. We discuss recent developments in systematic violence risk assessment, focusing on the Psycho-pathology Checklist (PCL) as a predictor, and examine data from key meta-analytic studies in the field.

HISTORY OF VIOLENCE PREDICTION

Predicting future risk of violent behaviour has a long and difficult history. Before 1966 relatively little attention was paid to how well clinicians assessed risk. The Baxstrom *v.* Herald (1966) ruling in the USA (which resulted in the release or transfer from maximum security hospitals of 966 patients to the community or to lower security) was a notable landmark in risk assessment history. Steadman & Coccoza (1974) reported on the 4-year outcomes of this cohort and found that only 20% had been reconvicted, the majority for non-violent offences. Throughout the 1970s several other studies reported in the literature fuelled the notions that clinicians had little expertise in predicting violent outcomes (e.g. Coccoza & Steadman, 1976; Thornberry & Jacoby, 1979).

Monahan (1984) reviewed these 'first generation' studies and concluded that "the upper bound level of accuracy that even the best risk assessment technology could achieve was of the order of 0.33". He reported that the best predictors of violence among the mentally disordered were the same demographic factors that predicted violence among non-disordered people, and that the poorest predictors were psychological factors such as diagnosis or personality traits. Subsequent studies, however, challenged these conclusions, particularly those demonstrating links between rates of violent offending and specific clinical diagnoses (e.g. Taylor, 1982; Binder & McNeil, 1988). The recent MacArthur Violence Risk Assessment Study (VRAS; Monahan *et al.*, 2000) also highlights the significance of clinical factors such as substance misuse and psychopathy as assessed by Hare's (1991) criteria, in the prediction of violent outcomes in non-forensic psychiatric patients discharged from hospital (see Steadman *et al.*, 1994; 1998).

GENERAL ISSUES IN VIOLENCE RISK PREDICTION

Clinical *v.* research perspectives

Clinicians have traditionally assessed violence risk on an individual basis, using a case formulation approach, i.e. 'unaided clinical judgement'. Until recently, however, research tended to focus on the accuracy of risk prediction variables in large, often heterogeneous, populations using relatively static actuarial predictors. These divergent approaches have resulted in debate over the merits of clinical *v.* actuarial approaches and their relevance to risk prediction for groups *v.* individuals. Furthermore, they have resulted in different perceptions about the relative contribution of clinical items in risk prediction scales in forensic and non-forensic settings. The clinical *v.* actuarial debate, however, has also led to the development of violence risk prediction instruments which adopt a combined approach and recognise the importance of both static actuarial variables and the clinical/risk management items that clinicians normally take into account in risk assessments of individuals. The latter approach appears to be a first step in bridging the gap between clinical and actuarial measures, and between group

and individual risk assessment approaches. In the following sections some of the key issues pertaining to clinical, actuarial and structured clinical assessments in violence risk prediction are discussed.

Approaches to violence risk prediction

Unaided clinical risk assessment

In clinical practice, assessments of the risk of dangerousness or violence in an individual are usually based solely on unaided clinical judgement. The unstructured clinical judgement approach to risk assessment has been criticised on a number of grounds, including low interrater reliability, low validity and a failure to specify the decision-making process (Monahan & Steadman, 1994; Webster *et al.*, 1997a), and inferior predictive validity compared to actuarial predictions (Meehl, 1954; Lidz *et al.*, 1993; Mossman, 1994). Others, however, consider that clinical approaches offer the advantages of flexibility and an emphasis on violence prevention (Snowden, 1997; Hart, 1998a). Buchanan (1999) also suggests that clinical approaches, if they focus on mechanisms through which violence occurs, may enhance the validity of risk assessment.

Clinicians may be better than was believed in the immediate aftermath of Baxstrom studies (Cocozza & Steadman, 1976). Gardner *et al.* (1996), for example, showed that while actuarial measures were better than clinical ratings, clinical ratings were better than chance. Studies also showed that the accuracy of prediction can be enhanced when clinicians consider the context in which violence occurs in their patients (Mulvey & Lidz, 1985). Recently, Fuller & Cowan (1999) showed that multi-disciplinary team consensus predictions of risk were comparable with actuarially based schedules over similar time-scales.

Actuarial methods

Actuarial methods allow assessors to make decisions based on data which can be coded in a predetermined manner (Meehl, 1954). Decisions are made according to rules, and focus on relatively small numbers of risk factors that are known, or are thought, to predict violence across settings and individuals. For diverse samples and contexts, these factors tend to be static (e.g. demographic variables). Actuarial approaches

undoubtedly improve the consistency of risk assessment, but Hart (1988a,b) argues that they tend to ignore individual variations in risk, overfocus on relatively static variables, fail to prioritise clinically relevant variables and minimise the role of professional judgement.

Despite these criticisms, actuarial risk assessment tools have been utilised for some time in US penal settings to help in making decisions about parole. Examples include the Base Expectancy Score (Gottfredson & Bonds, 1961), the Level of Supervision Inventory (Andrews, 1982), the Salient Factor Score (revised) (Hoffman, 1983), and the Statistical Information on Recidivism (SIR) scale (Nuffield, 1989). In the UK, similar measures have been developed to produce 'risk of reconviction' scores for prisoners before the parole board (Copas *et al.*, 1996).

Structured clinical judgement

Structured clinical judgement represents a composite of empirical knowledge and clinical/professional expertise. Webster *et al.* (1997a), who are the leading proponents of this model, argue that clinical violence risk prediction can be improved significantly if:

- (a) assessments are conducted using well-defined published schema;
- (b) agreement between assessors is good, through their training, knowledge and expertise;
- (c) prediction is for a defined type of violent behaviour over a set period;
- (d) violent acts are detectable and recorded;
- (e) all relevant information is available and substantiated;
- (f) actuarial estimates are adjusted only if there is sufficient justification.

Several instruments have been developed along these lines to assess risk of violence in clinical contexts. These include the Historical/Clinical/Risk Management 20-item (HCR-20) scale (Webster *et al.*, 1997b) the Spousal Assault Risk Assessment guide (Kropp *et al.*, 1995) and the Sexual Violence Risk (SVR-20) scale (Boer *et al.*, 1997) (see Douglas & Cox (1999) for an in-depth review of these instruments).

Hart (1998a,b) suggests that structured clinical instruments like the above promote systematic data collection based on sound scientific knowledge, yet allow flexibility in

the assessment process. He also argues that, unlike strict actuarial measures, they encourage clinicians to use professional discretion.

Violence risk prediction in clinical settings

A number of violence risk prediction tools have been developed and introduced into clinical settings in North America. Among these, the Dangerous Behaviour Rating Scale (DBRS: Menzies *et al.*, 1985a,b), the Violence Risk Appraisal Guide (VRAG: Harris *et al.*, 1993) and the HCR-20 (Webster *et al.*, 1997b) have received most attention. The last two instruments contain an item assessing psychopathy, based on the Psychopathy Checklist (revised) (PCL-R; Hare, 1991). The PCL-R itself, however, has also been shown to have reasonable predictive validity in determining future violence, and will also be discussed in some detail.

Before describing these tools and their predictive validity it may be useful to describe one of the more recent statistical measures which is frequently cited in the literature on the accuracy of violence risk prediction.

Statistical measures for assessing predictive accuracy

There are several measures available to evaluate the predictive accuracy of different tools in studies on violence risk prediction in large cohorts (see Appendix). Receiver operator characteristics (ROCs), which yield an area under the curve (AUC) measure, however, appear to be the preferred method, and much of the recent literature on predictive accuracy quotes ROC-AUC data. ROCs are particularly useful as they provide data which are fairly independent of the base rates of violence in a given population (Mossman, 1994). The ROC-AUC parameter, which can range from 0 to 1, provides information which is similar to that yielded by the more commonly used effect size estimate (such as Cohen's *d*; see Cohen, 1988; 1992) and can be used to compare accuracy between instruments. Figure 1 shows an example of a ROC curve. The straight line on the ROC curve corresponds to the line of no information, i.e. no better than random prediction (AUC=0.5). Instruments or clinicians which distinguish violent from non-violent patients with nearly perfect accuracy would have ROC-AUCs approaching 1.0. In general, Cohen's

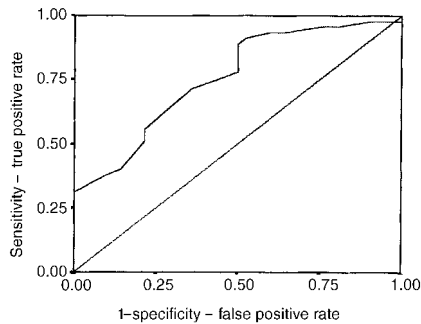


Fig. 1 Receiver operating characteristic curve (ROC) for PCL-SV. Based on unpublished data available from the first author upon request. Straight line=area under curve (AUC) 0.5 (no greater than chance prediction); curved line indicates PCL-SV potential decision threshold/cut-off scores; AUC for PCL-SV=0.76, $P < 0.001$.

$d > 0.50$ or ROC-AUCs > 0.75 are considered large effect sizes. ROC curves also give an indication of the trade-offs between specificity and sensitivity at different decision thresholds or cut-off scores on measures.

CLINICAL RISK ASSESSMENT TOOLS AND THEIR PREDICTIVE VALIDITY

The Dangerous Behaviour Rating Scale (DBRS)

The DBRS (Menzies *et al*, 1985a) was initially developed in conjunction with clinicians from a model devised by Megaree (1976). The item list comprised 18 ratings of personality, situation, lifestyle-related variables, and interview-specific factors believed to relate to risk. Reports on the predictive validity of the DBRS, however, indicate that it has met with little success. For example, in Menzies and Webster's (1995) 6-year follow-up study of 162 Canadian mentally disordered persons tracked across institutions and the community, comparatively little association was found between actuarial and clinical risk factors and follow-up violence outcome data. Professionals were no more accurate than non-clinical raters, and the DBRS items showed less predictive power than attributes such as age, violent history or employment status. The DBRS is now rarely used, and Webster *et al* (1997a) argue that the limitations of this instrument reflect the limited literature on which it was based.

The Violence Risk Appraisal Guide (VRAG)

The VRAG (Harris *et al*, 1993) incorporates 12 items, which are scored on the basis of a weighting procedure developed on a calibration sample of 618 males charged with severe violent crimes. The items are listed in Table 1. The variable with the heaviest weighting is the PCL-R score. Using ROCs, Rice & Harris (1995) analysed the data from several populations of offenders independent of the calibration sample, and found that the VRAG predicted violent recidivism with AUCs of 0.75, 0.74 and 0.74 for 3.5, 6 and 10 years, respectively. Using more restrictive definitions of violent recidivism, the relevant normalised ROC gave a mean AUC of 0.73. Later reports, however, suggest that the VRAG is less valuable in predicting violent sexual recidivism in paedophile sex offender populations (Rice & Harris, 1997). The VRAG has been criticised because of its reliance on relatively static factors, and Webster *et al* (1994) now recommend that it be supplemented with a clinical checklist to produce a 'violence prediction scheme'.

Psychopathy Checklist (Revised) (PCL-R)

The 20-item PCL-R (Hare, 1991), which is scored on a three-point scale, was originally devised as a research tool for operationalising psychopathy (see Table 1). Scores range from 0 to 40, with a cut-off of > 30 reflecting a prototypical psychopath. The PCL-R has been shown to have good psychometric properties (Cooke, 1998). It has a stable factor structure (Hare, 1991), in which factor 1 taps interpersonal/affective traits, while factor 2 reflects the behavioural components of psychopathy. Cooke & Michie (1998), however, have recently presented a three-factor model of psychopathy using confirmatory factor analytic procedures. A number of studies demonstrate its utility as a risk assessment tool, in identifying recidivists and predicting violence in North American forensic and prison samples (Hart, 1998a). As yet, there are few data on its predictive validity in European samples, although recent work by Grann *et al* (1999) suggests that the PCL-R scores were the best predictor of violent recidivism 2 years after release from containment in Swedish offenders with personality disorder (AUC=0.72).

The PCL-R has been supplemented by the 12-item screening version (PCL-SV; Hart *et al*, 1995) (Table 1). It has similar psychometric properties to the PCL-R, with scores ranging from 0–24 (cut-off at 18). The PCL-SV has been shown to have good predictive validity for institutional violence (Hill *et al*, 1996; Grann, 1998) and community violence (MacArthur study, Monahan *et al*, 2000).

As some psychopathy checklist items may be linked to outcome variables of interest (such as violence), researchers have used different methods to control for this potential confounder, including statistical control for past criminal activity or removing potentially confounding items from the checklist in the analysis.

The Historical/Clinical/Risk Management 20-item (HCR-20) scale

The HCR-20 (Webster *et al*, 1997b) contains 10 historical (H-10) items (two of which address the issue of personality dysfunction), five clinical (C-5) items, and five risk management (R-5) items (Table 1). It is scored in a similar manner to the PCL-R and shows good interrater reliability (Webster *et al*, 1997a). When the personality disorder variable is removed, H-10 items show significant correlations with on-ward violence (unpublished, 1996; details available from the first author upon request). In two studies, the H-10 items showed stronger correlations with violent outcome than the C-5 scales (see Douglas, 1996; unpublished, 1996); this may reflect the lack of inclusion of interview data in these retrospective studies.

Table 2 lists some key HCR-20 studies examining the predictive validity using ROC-AUC data. While the studies are limited to a small group of North American researchers, the data generally show 'better than chance' relationships between HCR-20 scores and violent outcomes. As yet, no studies have been published of the reliability and validity of this instrument in UK samples, although such work is in progress (details available from the first author upon request).

Comparison of actuarial risk scales

The PCL/PCL-R has generally been found to be superior to other classical actuarial risk scales on indices of recidivism or violent recidivism (Harris *et al*, 1993; Rice & Harris, 1995; Zamble & Palmer, 1996; see also

Table 1 Content lists for key risk prediction tools

PCL-R (Hare, 1991)	PCL-SV (Hart et al., 1995)	HCR-20 ¹ (V2) (Webster et al., 1997b)	VRAG (Harris et al., 1993)	MacArthur VRAS (Monahan et al., 2000)
				Predictors of community violence
				Iterative classificatory tree
1 Gibness/superficial charm (1)	1 Superficial (1)	H1 Previous violence	1 PCL-SV score	Psychopathy (PCL-SV > 12)
2 Grandiose sense of self-worth (1)	2 Grandiose (1)	H2 Young age at first violent incident	2 Elementary school maladjustment	Number of adult arrests
3 Need for stimulation/proneness to boredom (2)	3 Manipulative (1)	H3 Relationship instability	3 DSM-III diagnosis of personality disorder	Substance misuse
4 Pathological lying (1)	4 Lacks remorse (1)	H4 Employment problems	4 Age at index offence	Anger
5 Conning/manipulative (1)	5 Lacks empathy (1)	H5 Substance misuse problems	5 Lived with both parents to age 16	Father ever used drugs
6 Lack of remorse or guilt (1)	6 Does not accept responsibility (1)	H6 Major mental illness	6 Failure on prior conditional release	Father ever been arrested
7 Shallow affect (1)	7 Impulsive (2)	H7 Psychopathy (PCL-R/PCL-SV)	7 Non-violent offence score	Child abuse
8 Callous/lack of sympathy (1)	8 Poor behaviour controls (2)	H8 Early maladjustment	8 Marital status	Victim of child abuse
9 Parasitic lifestyle (2)	9 Lacks (goals) (2)	H9 Personality disorder	9 DSM-III diagnosis of schizophrenia	Recent violent behaviour
10 Poor behavioural controls (2)	10 Irresponsible (2)	H10 Prior supervision failure	10 Victim injury (index offence)	Violent fantasies
11 Promiscuous sexual behaviour (2)	11 Adolescent antisocial behaviour (2)	C1 Lack of insight	11 History of alcohol misuse	Anger reaction
12 Early behavioural problems (2)	12 Adult antisocial behaviour (2)	C2 Negative attitudes	12 Female victim (index offence)	Employed
13 Lack of realistic long-term goals (2)		C3 Active symptoms of major mental illness		Recent violence
14 Impulsivity (2)		C4 Impulsivity		Loss of conscience
15 Irresponsibility (2)		C5 Unresponsive to treatment		Parents fought
16 Failure to accept responsibility		R1 Plans lack feasibility		
17 Many short-term marital relationships		R2 Exposure to destabilisers		
18 Juvenile delinquency (2)		R3 Lack of personal support		
19 Revocation of conditional release (2)		R4 Non-compliance with remediation attempts		
20 Criminal versatility		R5 Stress		

1. H1-10 relate to history, C1-5 to clinical and R1-5 to risk.
 PCL-R, Psychopathy Checklist (Revised); PCL-SV, Psychopathy Checklist: Screening Version; HCR-20 (V2), Historical/Clinical/Risk Management 20-item scale, Version 2; VRAS, Violence Risk Assessment Study; VRAG, Violence Risk Appraisal Guide.
 (1), factor 1 loading; (2), factor 2 loading.

Hemphill *et al.*, 1998). Rice and Harris (1995) compared the SIR scale and VRAG, and found significantly better prediction rates with the VRAG, although the SIR scale (contrary to initial perceptions: Nuffield, 1982) also showed reasonable ROC-AUCs (0.69, 0.67 and 0.66 at 3.5-, 6- and 10-year follow-up). Zamble and Palmer (1996) compared the PCL-R, parole board decisions and the SIR scale, in 106 male offenders released from Canadian federal penitentiaries and found the PCL-R to be the most accurate at predicting reconviction or revocation of parole at a mean follow-up time of 30 months. Hemphill and Hare (1996) also compared the predictive validity of the PCL/PCL-R and several actuarial measures, and found that they performed similarly for general recidivism prediction, but that the PCL-R was significantly better for violent recidivism prediction.

Using ROCs, Grann (1998) compared the H-10 scale of the HCR-20 and VRAG in predicting reconviction for violence within 2 years of release, in a retrospective study of 293 violent offenders with personality disorders and 111 with schizophrenia. He reported that both scales performed better in the personality-disorder group but the H-10 did better than the VRAG in both groups of offenders. It is possible that historical/static variables may be relatively good predictors of violent recidivism in subjects with personality disorder, but clinical and risk management variables may be better predictors in populations with schizophrenia (Webster *et al.*, 1997a; Grann, 1998; Strand *et al.*, 1999).

General reviews and meta-analysis of studies of violence and recidivism

There have been four relatively recent meta-analytic studies of recidivism, including violent recidivism, and each differs in the studies included in it and the method of effect size determination.

Mossman (1994) extracted 58 data sets from 44 published studies dating from 1972 to 1993 on violence risk prediction, and examined prediction accuracy using ROCs. The studies included a broad range of subjects, settings, population sizes and clinical criteria for assessing violence, and Mossman acknowledges that conclusions can only be tentative. The median ROC-AUCs for all 58 data sets was 0.73, suggesting, overall, that clinicians were predicting violence more accurately than chance. However, short-term (1–7 days, AUC=0.68) predictions were no more accurate than long-term (>12 months, AUC=0.64). First-generation studies (before 1986) (AUC=0.74) were less accurate than second-generation studies (after 1986) (mean AUC=0.83), but the samples were extremely heterogeneous. Mossman suggests that clinicians were able to distinguish violent from non-violent patients with a “modest, better than chance level of accuracy”. Since this work was published, other reviews have concentrated on the issue of recidivism, particularly violent recidivism, which is generally perceived as a ‘harder’ outcome measure.

Bonta *et al.* (1998) conducted a meta-analysis of predictive longitudinal studies (1959–1995), to examine whether predictors of recidivism, including violent recidivism, for mentally disordered offenders were different from those for non-disordered offenders. Using 64 separate samples with 27 predictors for violent recidivism, they showed that criminal history variables were better predictors than clinical variables, using adjusted and transformed Pearson’s correlations to assess effect size (Z_r). For violent recidivism, criminal history variables had the largest effect size ($Z_r=0.15$, $P<0.001$), followed by personal demographics ($Z_r=0.12$, $P<0.001$), deviant lifestyle ($Z_r=0.08$, $P<0.001$) and clinical variables ($Z_r=-0.03$, $P<0.01$). A diagnosis of anti-social personality disorder was the most significant clinical predictor.

Role of the PCL/PCL-R and PCL-SV in risk prediction

The PCL-R and PCL-SV are currently believed by some of the most reliable tools for assessing personality constructs likely to be relevant to violent risk prediction (Hart, 1998b). For this reason the PCL-SV was included in the MacArthur VRAS, where it was shown to have reasonable predictive validity for community violence (Monahan *et al.*, 2000). Hemphill & Hare (1996) have also shown that the PCL/PCL-R, entered into a hierarchical multiple regression analysis with other demographic/clinical history variables,

Table 2 HCR-20 risk predictive validity studies

Author	Patient sample	Measures	Outcomes	Instrument	ROC-AUCs	Instrument	ROC-AUCs
Douglas <i>et al.</i> (1999)	Non-forensic psychiatric	Community violence	Violent crime	HCR-20	0.80	PCL-SV	0.78
			Any violence	HCR-20	0.76	PCL-SV	0.68
			Physical violence	HCR-20	0.76	PCL-SV	0.73
Nicholls <i>et al.</i> (1997)	Non-forensic psychiatric	Community violence (males)	Any violence	HCR-20	0.74		
			Violent arrest	HCR-20	0.78		
			Community violence (females)	Any violence	HCR-20	0.63	
Unpublished (1998a) ¹	Civil psychiatric patients	Community violence	Any violence	HCR-20	0.67	PCL-SV	0.65
			Violent arrest	HCR-20	0.75	PCL-SV	0.70
Grann (1998)	Forensic	Community violence	Violence in personality disorder & people with schizophrenia	H scale	0.71	VRAG	0.63
				H scale	0.66	VRAG	0.60
Unpublished (1997) ¹	Non-forensic psychiatric	In-patient violence	Physical & non-physical	H/C scales	0.57–0.65	PCL-SV	0.60–0.64
Unpublished (1998b) ¹	Non-forensic psychiatric	In-patient violence	Any type of aggression	H/C scales	0.63	PCL-SV	0.61
				H/C scales	0.68		

1. Details available from the first author upon request.

HCR-20, Historical/Clinical/Risk Management 20-item scale; PCL-SV, Psychopathy Checklist – Screening Version; ROC, receiver operator characteristics; AUC, area under curve.

adds significant incremental validity to the prediction of violence.

Meta-analytic studies using the PCL/PCL-R/PCL-SV in risk prediction

Salekin *et al* (1996) examined all 18 available (published and unpublished) studies using the PCL/PCL-R between 1974 and 1995 and conducted a meta-analytic

review, using an adaptation for effect size calculation from Rosenthal (1991). Separate analyses were conducted for violent recidivism. Despite the variation in cut-off scores on these instruments, Salekin *et al* (1996) reported moderate to strong effect sizes (Cohen's $d=0.55$ for criminality, $r=0.37$ and $d=0.79$ for violent recidivism; see Table 3). The largest effect sizes were reported in the study of institu-

tion violence by Hill *et al* (1996) using the PCL-SV. Although the study by Salekin *et al* (1996) included a small number of postdictive studies (comparing assessment measures with previous violence) which may have inflated their reported mean effect sizes, they found no significant difference between postdictive (0.75) and predictive (0.79) effect sizes, on a separate analysis.

Table 3 Studies utilising Psychopathy Checklists (PCLs) to predict violent behaviour

Study	Version	Sample	Validity	Outcome	Effect size (ES)
<i>Salekin et al (1996) review</i>					$r=0.37$ $d=0.80$
Forth <i>et al</i> (1990) ¹	PCL	Maximum security youth detention centre	Postdiction	Violent recidivism	$d=0.56$
Forth <i>et al</i> (1990) ¹	PCL	Maximum security youth detention centre	Prediction	Violent recidivism	$d=0.54$
Forth <i>et al</i> (1990) ¹	PCL	Maximum security youth detention centre	Postdiction	Institutional violence and/or aggression	$d=1.04$
Hare & McPherson (1984) ¹	PCL	Federal medium security prison	Postdiction	Violent recidivism	$d=0.54$
Harris <i>et al</i> (1991) ¹	PCI	Therapeutic community programme	Prediction	Violent recidivism	$d=0.93$
Heilbrun <i>et al</i> (1998) ¹	PCL	Forensic psychiatric patients	Prediction	Institutional violence and/or aggression	$d=0.63$
Hill <i>et al</i> (1996) ¹	PCL-SV	Forensic psychiatric patients	Prediction	Institutional violence and/or aggression	$d=1.92$
Kosson <i>et al</i> (1990) ¹	PCL-R	Maximum security youth detention centre	Postdiction	Violent recidivism	$d=0.42$
Miller <i>et al</i> (1994) ¹	PCL-R	Forensic treatment centre for sexual offenders	Postdiction	Violent recidivism	$d=1.18$
Quinsey <i>et al</i> (1995) ¹	PCL-R	Forensic treatment centre for sexual offenders	Prediction	Violent recidivism	$d=0.7$
Rice & Harris (1992) ¹	PCL	Forensic psychiatric patients	Prediction	Violent recidivism	$d=0.54$
Unpublished (1992) ³	PCL-R	Therapeutic community programme	Prediction	Violent recidivism	$d=0.72$
Rice <i>et al</i> (1990) ¹	PCL	Forensic treatment centre for sexual offenders	Prediction	Violent recidivism	$d=0.74$
Serin (1991)	PCL	Federal medium security prison	Postdiction	Violent recidivism	$d=0.74$
Serin & Amos (1995) ^{1,2}	PCL-R	Federal medium security prison	Prediction	Violent recidivism	$d=0.58$
<i>Hemphill et al (1998) review</i>					$r=0.27$ $d=0.56$
Hemphill <i>et al</i> (1992) ²	PCL	Criminal psychopaths	Prediction	Violent recidivism	$r=0.06$
Harris <i>et al</i> (1993) ²	PCL	Forensic psychiatric patients	Prediction	Violent recidivism	$r=0.34$
Heilbrun <i>et al</i> (1998) ²	PCL	Forensic psychiatric patients	Prediction	Violent recidivism	$r=0.16$
Ross <i>et al</i> (1992) ²	PCL	French prison parolees	Prediction	Violent recidivism	$r=0.17$
Serin & Amos (1995) ^{1,2}	PCL-R	Federal medium security prison	Prediction	Violent recidivism	$r=0.28$
<i>Studies outside North America</i>					
Grann <i>et al</i> (1999)	PCL-R	Forensic psychiatric evaluations (prison, hospital and probation)	Prediction	Violent recidivism	AUC of ROC at 2 years=0.75
Unpublished (1999) ³	PCL-SV	Medium security forensic	Prediction	In-patient violence	AUC of ROC at 3 months=0.75

1. Included in Salekin *et al* (1996). ES=mean PCL/PCL-R score for recidivists-non-recidivists/pooled s.d. of two groups (mean Cohen's $d=0.79$, $r=0.37$).

2. Included in Hemphill *et al* (1998). ES=raw PCL mean/recidivism correlation coefficient to Fisher Z s averaging these values, then converting back to r s (mean Cohen's $d=0.56$, $r=0.27$).

3. Further details available from the first author upon request.

SV, Screening Version; R, Revised; ROC, receiver operator characteristics, AUC, area under the curve.

Hemphill *et al* (1998) also conducted a meta-analysis of PCL/PCL-R studies in prediction of general/violent recidivism, but included only predictive studies and those with independent samples. The 1996 review by Salekin *et al* had included several same sample studies from the Oak Ridge group. Based on the five studies shown in Table 3 (1374 offenders) and more restrictive criteria, Hemphill *et al* (1998) reported a slightly lower mean effect size for violent recidivism ($r=0.27$, Cohen's $d=0.56$). Overall the predictive validity of the PCL-R is moderately high (Hart, 1998a).

PUTTING SYSTEMATIC RISK ASSESSMENT INTO PRACTICE

Gardner *et al* (1996) suggest that clinicians may be averse to actuarial or structured clinical prediction instruments because they are impractical and too costly, and the analyses too complex. They developed a 'regression tree' (i.e. structured sequences of yes/no answers that lead to classification of a case as high or low risk) and a two-stage screening process, which they showed was as accurate as traditional actuarial measures. Monahan *et al* (2000) also developed an Iterative Classification Tree (ICT) which successfully classified 77.6% of their sample as high or low risk, based on the variables shown in Table 2. ROCs for the ICT method were high (AUC=0.82). In higher risk cases, Serin and Amos (1995) suggest a three-stage decision tree, which includes an analysis of 'group base rates' of violence, 'individual base rate risk' and 'risk management variables'.

Decision or classification trees appear to be a useful means of streamlining violence risk assessments in large populations with relatively low base rates of violence. In smaller samples of high-risk patients or offenders, however, more in-depth batteries of relevant tools such as the PCL-R and HCR-20 will be required to assess future risk of violent recidivism.

SUMMARY

This review indicates that structured clinical judgement and systematic risk assessment scales should be used cautiously and judiciously. The assessment tools chosen, and how to interpret the scores, will largely be influenced by the populations or settings and the questions we

want answered. The MacArthur project group have developed a classificatory tree method for assessing risk in community samples. For clinicians working in forensic or penal settings, much can be learned from the studies demonstrating the predictive accuracy of tools such as the PCL-R and HCR-20. Future British studies should aim to establish the validity of North American risk assessment tools in a range of populations and settings. Efforts should also be made to enhance the predictive validity of these tools by the addition of physiological measures and assessments of neurocognitive function and how individuals process emotional information. Violence prediction will never be entirely accurate, given that violence itself is a complex concept. Clinicians need to be aware of the benefits and limitations of current assessment tools and how scores on these measures might be used or interpreted by other agencies (see Hare (1998) for a commentary on the use and misuse of the PCL-R).

APPENDIX – TERMINOLOGY

True positive, TP=predicted risk, outcome violent

False positive, FP=predicted risk, outcome not violent

False negative, FN=predicted no risk, outcome violent

True negative, TN=predicted no risk, outcome not violent

Base rate, BR=(TP+FN)/(TP+FP+FN+TN)=(proportion of violent individuals in a population)

Selection ratio, SR=(TP+FP)/(TP+FP+FN+TN)
=(cut-off scores used to classify individuals as violent)

Correct fraction, CF=(TP+TN)/(TP+FP+FN+TN)

Sensitivity=true positive rate, TPR=TP/(TP+FN)

Specificity=true negative rate, TNR=TN/(TN+FP)

Positive predictive power=Proportion of individuals designated a risk who in fact are a risk

Negative predictive power=Proportion of individuals identified as low risk and who in fact are low risk

False positive rate, FPR=(1-specificity)=FP/(FP+TN)

Risk ratio=TPR/FPR

Odds ratio=(TP.TN)/(FP.FN)=odds that person predicted to fail will do so/odds a person not predicted to fail will do so

Relative improvement over chance, RIOC=CF - ((BR)(SR)+(1-BR)(1-SR))

REFERENCES

- Andrews, D. A. (1982) *The Level of Supervision Inventory (LSI)*. Toronto, Canada: Ontario Ministry of Correctional Services.
- Baxstrom v. Herald (1966) 383 U.S. 107.
- Binder, R. & McNeil, D. (1988) Effects of diagnosis and context of dangerousness. *American Journal of Psychiatry*, **145**, 728–732.
- Boer, D. P., Hart, S. D., Kropp, P. R., et al (1997) *Manual for the Sexual Violence Risk-20: Professional Guidelines for Assessing Risk of Sexual Violence*. Vancouver, B.C.: British Columbia Institute on Family Violence.
- Bonta, J., Hanson, K., Law, M. (1998) The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, **123**, 123–142.
- Buchanan, A. (1999) Risk and dangerousness. *Psychological Medicine*, **29**, 465–473.
- Cocozza, J. J. & Steadman, H. J. (1976) The failure of psychiatric predictions of dangerousness: Clear and convincing evidence. *Rutgers Law Review*, **29**, 1084–1101.
- Cohen, J. (1988) *Statistical Power Analysis for the Behavioural Sciences*. New York: Academic Press.
- (1992) A power primer. *Psychological Bulletin*, **112**, 155–159.
- Cooke, D. J. (1998) Cross-cultural aspects of psychopathy. In *Psychopathy: Theory, Research and Implications for Society* (eds D. Cooke, A. E. Forth & R. D. Hare). Dordrecht, The Netherlands: Kluwer, pp. 13–47.
- & Mitche, C. (1998) Towards a hierarchical model of psychopathy. Unpublished manuscript.
- Copas, J., Marshall, P. & Tarling, R. (1996) Predicting re-offending for discretionary conditional release: *Home Office Research Study 150*. Home Office: London.
- Douglas, K. S. (1996) *Assessing the risk of violence in civil psychiatric outpatients: The predictive validity of the HCR-20 risk assessment scheme*. Unpublished master's thesis, Simon Fraser University, Burnaby, British Columbia.
- & Cox, D. N. (1999) Violence risk assessment: Science and practice. *Legal and Criminological Psychology*, **4**, 149–184.
- , Ogloff, J. R. P., Nicholls, T. L., et al (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*, in press.
- Forth, A. E., Hart, S. D. & Hare, R. D. (1990) Assessment of psychopathy in male young offenders. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, **2**, 342–344.
- Fuller, J. & Cowan, J. (1999) Risk assessment in a multi-disciplinary forensic setting: Clinical judgement revisited. *Journal of Forensic Psychiatry*, **10**, 276–289.
- Gardner, W., Lidz, C. W., Mulvey, E. P., et al (1996) A comparison of actuarial methods for identifying repetitively violent patients with mental illnesses. *Law and Human Behaviour*, **20**, 35–48.
- Gottfredson, D. M. & Bonds, J. A. (1961) *A manual for intake base expectancy scoring*. San Francisco, CA: California Department of Corrections, Research Division.
- Grann, M. (1998) *Personality Disorder and Violent Criminality: A Follow-Up Study with Special Reference to Psychopathy and Risk Assessment*. Stockholm: Karolinska Institute, Department of Clinical Neuroscience and Family Medicine.

—, **Langstrom, N., Tengstrom, A., et al (1999)** Psychopathy (PCL-R) predicts violent recidivism among criminal offenders with personality disorders in Sweden. *Law and Human Behaviour*, **23**, 205–218.

Hare, R. D. (1991) *Manual for the Hare Psychopathy Checklist-Revised*. Toronto, Canada: Multi-Health Systems.

— (1998) The Hare PCL-R: Some issues concerning its use and misuse. *Legal and Criminological Psychology*, **3**, 99–119.

— & **McPherson, L. M. (1984)** Violent and aggressive behaviour by criminal psychopaths. *International Journal of Law and Psychiatry*, **7**, 35–50.

—, **Harpur, T. J., Hakstian, A. R., et al (1990)** The Revised Psychopathy Checklist: Reliability and factor structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, **2**, 338–341.

Harris, G. T., Rice, M. E. & Cormier, C. A. (1991) Psychopathy and violent recidivism. *Law and Human Behaviour*, **15**, 625–637.

—, — & **Quinsey, V. L. (1993)** Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behaviour*, **20**, 315–335.

Hart, S. D. (1998a) Psychopathy and risk for violence. In *Psychopathy: Theory, Research and Implications for Society* (eds D. Cooke, A. E. Forth & R. D. Hare), pp. 355–375. Dordrecht, The Netherlands: Kluwer.

— (1998b) The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology*, **3**, 121–137.

—, **Cox, D. N. & Hare, R. D. (1995)** *The Hare Psychopathy Checklist – Revised Screening Version (PCL:SV)*. Toronto: Multi-Health Systems.

Heilbrun, K., Hart, S. D., Hare, R. D., et al (1998) Inpatient and post-discharge aggression in mentally disordered offenders. *Journal of Interpersonal Violence*, **13**, 514–527.

Hemphill, J. F. & Hare, R. D. (1996) Psychopathy Checklist factor scores and recidivism. In *International Perspectives on Psychopathy* (eds D. J. Cooke, A. E. Forth, J. Newman, et al), pp. 68–73. Leicester: British Psychological Society.

—, **Hare, R. D. & Wong, S. (1998)** Psychopathy and Recidivism: A Review. *Legal and Criminological Psychology*, **3**, 139–170.

Hill, C. D., Rogers, R. & Bickford, M. E. (1996) Predicting aggressive and socially disruptive behaviour in a maximum security forensic hospital. *Journal of Forensic Sciences*, **41**, 56–59.

Hoffman, P. B. (1983) Screening for risk: A revised Salient Factor Score (SFS:81). *Journal of Criminal Justice*, **1**, 539–547.

Kosson, D. S., Smith, S. S. & Newman, J. P. (1990) Evaluating the construct validity of psychopathy in black and white male inmates: Three preliminary studies. *Journal of Abnormal Psychology*, **99**, 250–259.

Kropp, P. R., Hart, S. D., Webster, C. W., et al (1995) *Manual for the Spousal Assault Risk Assessment Guide, 2nd ed.* Vancouver, BC: British Columbia Institute on Family Violence.

Lidz, C. W., Mulvey, E. P. & Gardner, W. (1993) The accuracy of predictions of violence to others. *Journal of the American Medical Association*, **269**, 1007–1111.

Meehl, P. E. (1954) *Clinical versus Statistical Prediction*. Minneapolis, MN: University of Minnesota Press.

Megaree, E. I. (1976) The prediction of dangerous behaviour. *Criminal Justice and Behaviour*, **3**, 3–22.

CLINICAL IMPLICATIONS

■ Structured/systematic approaches to violence risk prediction provide a more accurate and transparent record of the risk factors considered and the rationale behind decisions taken.

■ Risk assessment batteries need to be streamlined and adapted to suit the population under study and the key questions asked.

■ The Psychopathy Checklist and its derivatives appear to be significant predictors of violence in forensic and non-forensic settings.

LIMITATIONS

■ There are a limited number of studies on the reliability and validity of published risk assessment tools outside the centres in which they were developed.

■ The literature on accuracy of violence risk is predominately postdictive rather than predictive, and much needs to be done to improve current violence prediction accuracy, using prospective study designs.

■ The lack of uniformity in the statistical procedures used to assess predictive accuracy, and the variation in choice of cut-off scores on risk prediction tools, make comparisons between studies difficult. The reporting of receiver operator characteristic data should improve this situation.

MAIREAD DOLAN, MRCPsych, MICHAEL DOYLE, MSc, Edenfield Centre, Mental Health Services of Salford

Correspondence: Dr Mairead Dolan, Edenfield Centre, Mental Health Services of Salford, Bury New Road, Prestwich, Manchester M25 3BL

(First received 26 October 1999, final revision 12 April 2000, accepted 13 April 2000)

Menzies, R. J., Webster, C. D. & Sepejak, D. S. (1985a) The dimensions of dangerousness: Evaluating the accuracy of psychometric predictions of violence among forensic patients. *Law and Human Behaviour*, **9**, 35–56.

—, — & — (1985b) Hitting the forensic sound barrier: Predictions of dangerousness in a pre-trial psychiatric clinic. In *Dangerousness: Probability and prediction, psychiatry and public policy* (eds C. D. Webster, M. H. Ben-Aron & S. J. Hucker), pp. 115–143. New York: Cambridge University Press.

— & — (1995) Construction and validation of risk assessments in a six year follow-up of forensic patients: A tridimensional analysis. *Journal of Consulting and Clinical Psychology*, **63**, 766–778.

Monahan, J. (1984) The prediction of violent behaviour: Toward a second generation of theory and policy. *American Journal of Psychiatry*, **141**, 10–15.

— (1992) Mental disorder and violent behaviour. *American Psychologist*, **47**, 511–521.

— & **Steadman, H. J. (eds) (1994)** *Violence and Mental Disorder: Developments in Risk Assessment*. Chicago, IL: University of Chicago Press.

—, —, **Appelbaum, P. S., et al (2000)** Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry*, **176**, 312–320.

Mossman, D. (1994) Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, **62**, 783–792.

Mulvey, E. P. & Lidz, C. W. (1985) Back to basics: A critical analysis of dangerousness research in a new legal environment. *Law and Human Behaviour*, **9**, 209–219.

Nicholls, T. L., Douglas, K. & Ogloff, J. R. P. (1997) Risk assessments with female psychiatric patients: utility of the HCR-20 and PCL:SV. *Canadian Psychology*, **38**, 111–112.

Nuffield, J. (1982) Parole decision making in Canada: Research towards decision guidelines. Ottawa, Canada, Communication Division, Solicitor General of Canada.

— (1989) The 'SIR Scale': Some reflections on its applications. *Forum on Corrections Research*, **1**, 19–22.

Quinsey, V. L., Rice, M. E. & Harris, G. T. (1995) Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence*, **10**, 85–105.

Reed, J. (1997) Risk assessment and clinical risk management: the lessons from recent inquiries. *British Journal of Psychiatry*, **170** (suppl. 32), 4–7.

Rice, M. E., Harris, G. T. & Quinsey, V. L. (1990) A follow-up of rapists assessed in a maximum security psychiatric facility. *Journal of Interpersonal Violence*, **5**, 435–448.

- & — (1992) A comparison of criminal recidivism among schizophrenic and non-schizophrenic offenders. *International Journal of Law and Psychiatry*, **15**, 397–408.
- & — (1995) Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology*, **53**, 737–748.
- & — (1997) Cross validation and extension of the Violence Risk Appraisal Guide for child molesters and rapists. *Law and Human Behaviour*, **21**, 231–241.
- Rosenthal, R. (1991)** Meta-analysis: a review. *Psychosomatic Medicine*, **53**, 247–271.
- Salekin, R., Rogers, R. & Sewell, K. (1996)** A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist–Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice*, **3**, 203–215.
- Serin, R. C. (1991)** Psychopathy and violence in criminals. *Journal of Interpersonal Violence*, **6**, 423–431.
- & **Amos, N. L. (1995)** The role of psychopathy in the assessment of dangerousness. *International Journal of Law and Psychiatry*, **18**, 231–238.
- Snowden, P. (1997)** Practical aspects of clinical risk assessment and management. *British Journal of Psychiatry*, **170** (suppl. 32), 32–34.
- Steadman, H. J. & Cocozza, J. J. (1974)** *Careers of the Criminally Insane: Excessive Social Control of Deviance*. Lexington, MA: Lexington Books.
- , **Monahan, J., Appelbaum, P. S., et al (1994)** Designing a new generation of risk assessment research. In *Violence and Mental Disorder: Developments in Risk Assessment* (eds J. Monahan & H. J. Steadman), pp. 297–318. Chicago, IL: University of Chicago Press.
- , **Mulvey, E. P., Monahan, J., et al (1998)** Violence by people discharged from acute psychiatric inpatient facilities and by others on the same neighbourhood. *Archives of General Psychiatry*, **55**, 393–401.
- Strand, S., Belfrage, H., Fransson, G., et al (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.
- Taylor, P. (1982)** Schizophrenia and violence. In *Abnormal Offenders, Delinquency, and the Criminal Justice System* (eds J. Gunn & D. P. Farrington), pp. 269–284. New York: Wiley.
- Thornberry, T. P. & Jacoby, J. E. (1979)** *The Criminally Insane: A Community Follow-up of Mentally Ill Offenders*. Chicago, IL: University of Chicago Press.
- Webster, C. D., Harris, G. T., Rice, M. E., et al (1994)** *The Violence Prediction Scheme: Assessing Dangerousness in High Risk Men*. Toronto: University of Toronto, Centre of Criminology.
- , **Douglas, K. S., Eaves, D., et al (1997a)** Assessing risk of violence to others. In *Impulsivity: Theory, Assessment and Treatment* (eds C. D. Webster & M. A. Jackson). New York: Guilford Press.
- , —, —, **et al (1997b)** *HCR-20: Assessing risk of violence (version 2)*. Vancouver: Mental Health Law & Policy Institute, Simon Fraser University.
- Zamble, E. & Palmer, W. (1996)** Prediction of recidivism using psychopathy and other psychologically meaningful variables. In *International Perspectives on Psychopathy* (eds D. J. Cooke, A. E. Forth, J. Newman, et al), pp. 153–156. Leicester: British Psychological Society.